SVENSKA GEOFYSISKA FÖRENINGEN

VOLUME 1, NUMBER 3 **Tellus** AUGUST 1949

A QUARTERLY JOURNAL OF GEOPHYSICS

Isostasy and its Meaning

By B. GUTENBERG, California Institute of Technology[1]

(Manuscript received June 25, 1949)

Abstract

The theory of isostasy supposes that, in regions which have not been disturbed recently, each vertical column of the earth's crust with a certain minimum radius and extending to a depth of about 100 km has approximately the same mass. To find the deviation from this approximation in a given region, the density must be assumed as a function of depth. Such assumptions used at present for calculations are discussed critically. The resulting errors are greater than it is normally beleaved; errors in the calculated isostatic gravity anomalies exceeding ten milligals must be expected in certain regions. Systematic errors result from the usual assumption in routine calculations that the mean density in the earth's crustal layers under the bottom of the Pacific and in the continental areas is the same, and that in both the difference between the density of the layers above about 30 km and the layers below this depth is 0.6. The processes producing and maintaining isostatic equilibrium are discussed.

In the theory of isostasy it is assumed that in regions that have not been disturbed recently each vertical column of the earth's crust with a given radius (at least, say, 10 km) and extending down to a sufficient specified depth (apparently at least 60 km), has approximately the same mass, regardless of the surface condition (continental or oceanic) or of the surface elevation of the region. Since this hypothesis is an approximation, the question cannot be raised whether the theory of isostasy is true or false, but how good the approximation is in a given region.

In addition to the use of deflections of the vertical, for which too few observations are available, there are two ways to approach this question. The first is, to use observations of the density and distribution of rocks in the earth's crust and of gravity at the earth's surface and to calculate the residuals (the so-called gravity anomalies) against an assumed equilibrium condition; the second is to consider the processes involved in establishing isostasy. Neither provides an answer with the desired precision, since each requires certain assumptions which are not as well founded as is commonly believed.

For the first approach it is necessary to observe gravity at as large a number of points on the earth's surface as possible, and to calculate under various assumptions regarding the structure of the earth's crust their deviation from gravity values theoretically to be ex-

[1] Contribution No. 504, California Institute of Technology, Division of the Geological Sciences, Pasadena, California.

pected, if hydrostatic equilibrium exists at an assumed depth. Thus, these calculations must be based on hypotheses regarding the density in the various crustal layers as well as on the thickness of these layers. Two fundamental hypotheses have been used which are based on two extreme assumptions (for historical references, see e.g. BOWIE 1931). For the first, PRATT supposed that all crustal columns begin at a discontinuity which has the same depth for the whole earth, that within each such column the density is constant, and that the differences in elevations of the earth's crust are compensated by different densities in the various crustal columns. The assumed discontinuity is called "depth of compensation". The product of the density d of the column and the surface elevation h above the depth of compensation is the same everywhere:

$$dh = \text{constant} \qquad (1)$$

The other extreme assumption is that of AIRY who assumed that all blocks near the surface have the same density and "float" in the deeper material, like icebergs are floating in water. If we denote by d_1 the density of the deeper material in which the crust is floating, by d_2 the density of the floating material, by h_1 the distance of the bottom of the floating material from an arbitrary depth, which must be at or below the level of the deepest part of the floating material, and by h_2 the thickness of the floating material, then the AIRY hypothesis can be expressed by the following equation:

$$d_1 h_1 + d_2 h_2 = \text{constant} \qquad (2)$$

AIRY assumed that d_1 and d_2 do not depend on the locality under consideration but locally different values of d_2 are indicated by the observations (Table I) and are to be expected for geophysical reasons.

In addition to his idea concerning the mechanism by which isostasy is maintained, (but not as a prerequisite to it), PRATT assumed that the difference in density in the various parts of the earth's crust depends on the amount of contraction at the time when the material solidified, thus explaining the smaller density of high mountain areas. This idea of PRATT is generally considered as incorrect. On the other hand, in the AIRY theory it must be assumed that there is sufficient "plasticity" somewhere in the outer part of the earth to make "isostatic" adjustment possible.

There can be little doubt that neither the PRATT nor the AIRY hypothesis is completely accurate. In most regions the actual conditions are probably better approximated by the AIRY hypothesis than by PRATT's. (See GUTENBERG 1927 and HEISKANEN 1936). The condition for isostasy is approximately given by

$$\int_0^h d\, dz = \text{constant} \qquad (3)$$

regardless of the locality; the density d is a function of the depth, different from region to region, z is the elevation above a level at which approximately hydrostatic equilibrium can be assumed, say about 100 km below sea level, and h is the value of z at the earth's surface.

There have been various ways by which the mechanism of isostasy and the deviations of the actual from assumed ideal conditions have been investigated. In the early attempts (see e. g. HEISKANEN 1936) it had been believed that the best approximation to the actual conditions could be found by assuming a variety of density distributions and calculating the resulting gravity anomalies; that assumption was considered the most likely, which gave the smallest anomalies. More recently (HEISKANEN 1948) the assumption has been considered best which gives the least effect of elevation on the residuals in a given area. Finally, HEISKANEN (1936) calculated residuals assuming the probable thickness of the layers as they have been found from seismological evidence and combining them with probable values of the density in these layers. This type of reduction, which is occasionally referred to as "HEISKANEN method", certainly is preferable to any other. Unfortunately, it requires a large amount of work for the calculations.

Recently the tendency prevails to use still more uniform assumptions for the whole earth and to apply them to routine reductions of the fast increasing body of data, rather than make for each region such individual assumptions which seem to be most probable for the local tectonic structure down to a depth of at least 60 km. Thus, it has been proposed recently (HEISKANEN 1948 a) to use not only the

same mean densities of 2.67 for the upper layers, and of 3.27 for the "sima" regardless of the crustal structure—continental, Atlantic or Pacific—but, in addition, to suppose everywhere the same "thickness of the crust" of 30 km (corresponding to a region where the surface is at sea level). If this suggestion is carried out, a faster calculation of "isostatic gravity anomalies" will be possible, and outlines of areas with large anomalies will become available more speedily; however, increased errors in the calculated anomalies must be expected. Conclusions based on differences of the order of 10 milligals, especially in comparing values in oceanic areas with those in continental regions must be expected to become even more doubtful than at present; unnecessary great uncertainties in the calculated isostatic gravity anomalies result from supposing everywhere densities of 2.67 and 3.27 in the "floating" and "supporting" layers respectively, regardless of the tectonic structure in the region involved. There is little doubt that the density of the surface layers in the continents is noticeably smaller than that of the corresponding crustal layers in the oceans. Data, according to WASHINGTON (1922), are given in Table I.

Table I

Average densities according to WASHINGTON *a* without, *b* with water content, *n* number of analyses.

Region	Average density		*n*
	a	*b*	
North America ..	2.78	2.73	1,709
South America ..	2.78	2.73	138
Europe	2.79	2.73	1,985
Africa	2.78	2.72	223
Australia	2.81	2.75	287
Atlantic	2.89	2.81	56
Pacific	3.09	3.01	72

Table I shows, in addition, that in all probability, even in the continents, the density assumed for the crustal layers is too small, considering the fact that values assumed for the gravity calculations have to represent not only the density of the uppermost surface layers, but also of the deeper more basic and heavier layers down to the material in which the upper layers are "floating". Their approximate thickness is given in the last column of Table II.

Table II

Average thickness of layers in the earth's crust with a given velocity of longitudinal earthquake waves (all data approximate with relatively large local variations).

Region	Thickness of layers, including sediments, in km		
	With velocity of longitudinal waves of		Both combined
	$5^1/_2$—$6^1/_4$ km/sec.	$6^1/_4$—$7^3/_4$ km/sec.	
Northwest Europe	27	small?	27
Black Forest	16	16	32
Central Alps	30	25	55
Central Asia	30	20	50
Southern California coastal.....	18	18	36
Central California coastal.....	10	22	32
Sierra Nevada ...	20	40	60
Eastern North America	10	30	40
New Zealand....	12	18	30
Atlantic Ocean ..	sediments?	20?	20?
Pacific Ocean....	sediments?	small?	small?

The density of rocks in these layers, as determined from laboratory samples under pressure of 1 atmosphere, varies from about 2.65 in granites, 2.7 in granodiorites and also in the Basement Complex in Finland, to 3.0 in gabbro, which by many is considered characteristic for the deeper layers of the "floating" material. The densities of rocks which are possibly characteristic of the layer "supporting" the crust, such as peridotite, pyroxenite and dunite, are about 3.2 to 3.3. Thus, in the continental areas, the difference in density between the "floating" and "supporting" material may well be only $^2/_3$ of the difference of 0.6, which is assumed for practically all gravity calculations, and still less in areas with "Pacific structure". The effect of pressure on the density can be assumed to be the same for both layers, as a first approximation.

It does not seem unlikely that the assumption of a difference in density of the two layers of 0.6 throughout the world results in systematic errors which affect especially calculations of gravity anomalies for the oceans. This may well be one of the reasons for the finding from gravity observations that the equator can be represented in second approximation by an ellipse with the short

and long axes pointing towards the continental and oceanic areas of the earth's crust respectively and differing by roughly 300 meters.

As a consequence of the inaccuries of the assumptions, the calculation of isostatic gravity anomalies involves errors which are far greater than the errors in the observations. Gravity observations frequently are accurate to better than 1 milligal. On the other hand, under favorable circumstances the calculated gravity anomalies may be uncertain by 10 milligals or more, especially in mountain areas. Table III gives selected data after HEISKANEN (1936). Still greater differences are to be found in his tabulations for more extreme values of the supposed depth of compensation; in addition, the anomalies calculated by the AIRY method correspond to the same mean density of the surface layers everywhere, and other assumptions will increase the differences between the results.

Table III

Isostatic gravity anomalies Δg in milligal after HEISKANEN (1936).

a) **Using PRATT'S reduction method, depth of compensation T km.**
b) **AIRY'S method, "thickness of the crust" D km.**
c) **HEISKANEN'S method, using geophysically determined data on the crustal layers.**

Region	Method of reduction								
	a		*a*		*b*		*b*		*c*
	Δg	T	Δg	T	Δg	D	Δg	D	Δg
87 stations in plains, USA	+ 8	85.3	+ 7	113.7	+ 6	60	+ 4	40	+ 4
20 mountain stations, USA	+ 18	85.3	+ 9	113.7	— 10	77	+ 11	40	+ 6
27 coastal stations, USA	+ 3	85.3	— 1	113.7	— 4	77	+ 4	40	— 1
23 coastal stations, Pacific, USA			— 20	113.7	— 20	60	— 18	40	— 15
12 stations, Caucasus	+ 47	156.3	+ 61	113.7	+ 28	77	+ 38	64	
10 stations, Alps	+ 31	85.3	+ 18	113.7	— 14	77	+ 19	40	

The belts of large negative and large positive anomalies stand out regardless of the method of reduction under reasonable assumptions. These belts of large gravity anomalies coincide almost everywhere with belts of earthquakes (GUTENBERG and RICHTER 1949). There is no doubt that they are actually regions in which tectonic processes are going on, disturbing continuously the isostatic equilibrium. There is some indication from the post-glacial uplift that in the earth's crust the relaxation time, that is the time during which a given deviation from equilibrium is reduced to $1/e$ by flow processes, is of the order of 10,000 years. No exact figure can be given for specific layers, especially since the effect of the crustal layers on this process is not known well enough. However, since deep focus earthquakes occur down to a depth of about 700 km, strain must be able to accumulate down to this depth; and the relaxation time in these layers must be expected to be at least of the order of 100 years. It may well be that at a depth of about 700 km, the relaxation time falls below the critical limit at which flow processes become so fast that they prevent accumulation of strain large enough to produce a break.

The speed of flow-processes is frequently characterized by the "coefficient of viscosity" η, defined by

$$\eta = \mu\tau \qquad (4)$$

where μ is the coefficient of rigidity (about 5×10^{11} dynes/cm^2 in the layers under consideration) and τ the time of relaxation. The coefficient of viscosity corresponding to a time of relaxation of 10,000 and 100 years is consequently of the order of 10^{23} and 10^{21} poises respectively. It cannot be appreciably below the latter value at depths down to

700 km, if the hypotheses involved in the preceding calculations are correct.

The second method for investigating the mechanism of isostasy is to consider the structures and the processes involved. In Table II a summary is given of the approximate thickness of layers in several regions from seismic data published by various seismologists. On the other hand, results found from isostatic calculations (HEISKANEN 1948 a, 1948 b) indicate that the "thickness of the crust" for a region with the surface at sea level is in general about 30 km and at greater depth under mountain ranges. The combined results can be interpreted as meaning that down to a depth of at least 60 km there are relatively large differences in density between the various parts of the earth. This does not mean necessarily that contemporary flow processes within this depth range produce approximate equilibrium. At present, noticeable flow processes may be limited to distinctly greater depths. It must be considered that the speed of flow-processes in a given region was probably appreciable higher at times when mountains were forming there under much greater stresses and higher temperatures than today, and that probably the "roots of mountains" originated simultaneously with the mountains.

Unfortunately, our theoretical knowledge of flow processes in the interior of the earth is practically nil. At present there is no theory available which applies to such processes. All that is known are results of a relatively few experiments on rock under high pressures or temperatures (rarely both), which have led to empirical descriptions of such processes. The "constants" involved are functions of time and a variety of physical conditions. Processes at the surface of the earth, which are rather slow compared with the time of relaxation of roughly 10,000 years, such as sedimentation and erosion, do not lead to large gravity anomalies or deviations from equilibrium. Where processes have a much higher speed, gravity anomalies must be expected, and vice-versa. If these processes stop, as, for example, in melting of ice or accumulation of ice in glaciated areas, these areas gradually approach hydrostatic equilibrium again.

The earth's crust as a whole is not in hydrostatic equilibrium. However, there is no doubt that there are large tectonically quiet areas in which the material at a depth of about 100 km is not very far from hydrostatic equilibrium. The combination of these facts is what we call "isostasy". One of our major aims must be to find deviations from hydrostatic equilibrium in the various regions of the earth with much higher accuracy than at present.

REFERENCES

BOWIE, W., 1931: Isostasy. Physics of the earth, vol. 2, Nat. Res. Council, Washington, D. C., No. 78, pp. 103—115.

GUTENBERG, B., 1927: Die Bedeutung der Isostasie. *Gerlands Beitr. z. Geophysik*, **16**: pp. 396—403.

GUTENBERG, B. and RICHTER, C. F., 1949: Seismicity of the Earth and Associated Phenomena. Princeton Press.

HEISKANEN, W., 1936: Das Problem der Isostasie. *Handbuch der Geophysik*, vol. 1, pp. 878—951. Berlin.

HEISKANEN. W., 1948 a: Report on Isostasy. Intern. Union of Geophys. and Geod., Oslo.

——, 1948 b: Lectures at Pasadena, December 1948. See also: On the interpretation of the gravity anomalies. *Geol. Soc. Amer. Bull.*, (in press).

MEINESZ, VENING F. A., 1933: Ergebnisse der Schwerkraftbeobachtungen auf dem Meere in den Jahren 1923—1932. *Ergebnisse der Kosmischen Physik*, **2**: pp. 153—212.

WASHINGTON, H., 1922: Isostasy and rock density. *Geol. Soc. Amer. Bull*, **23**: pp. 372—410.

Zur Theorie des Wärmehaushaltes strahlungsdurchlässiger Medien

Von H. REUTER

Zentralanstalt für Meteorologie und Geodynamik, Wien

(Manuscript received June 7, 1949)

Abstract

In view of the fact that large areas of the earth's surface are temporarily or permanently covered by a medium pervious to radiation, i.e. water, snow or ice, attention is directed to that special feature of the heat economy of these media having to do with radiation balance and heat transfer. Primarily, the purpose of the theory is to find an explanation for the observed distribution of temperature within these media. A first attempt to solve this problem was previously made by the author (REUTER 1948 a) in a study of the heat economy of snow cover. The essential basis of the present as well as the earlier treatment lies in the assumption that snow, ice and water, while pervious to insolation, are nevertheless impervious to outgoing radiation. The present paper treats only the simple case of constant incoming penetrating radiation and at the same time a constant effective outgoing radiation from the surface of the medium. One of the principal results is that the temperature maximum produced during the heating process does not appear at the surface, but rather at some depth below. Although the case of water is not treated in this report, the results obtained for ice may be applied, with slight modification, to the case of calm pure water in view of the use of the same coefficients of absorption for both water and ice.

Einleitung

Die qualitativen und quantitativen Unterschiede im Wärmehaushalt der verschiedenen die Erdoberfläche bildenden Medien stellen infolge ihrer verschiedenen Strahlungs- und Wärmeleitungseigenschaften einen Gegenstand erhöhten Interesses dar. Die Tatsache, das weite Gebiete der Erdoberfläche dauernd oder zeitweise mit strahlungsdurchlässigen Medien (Wasser, Schnee, Eis) bedeckt sind, lenkt die Aufmerksamkeit auf die Unterschiede in der Theorie des Wärmehaushaltes strahlungsdurchlässiger Medien gegenüber strahlungsundurchlässigen. Es handelt sich hier vor allem um die folgende Frage: Wie kommt die beobachtete Temperaturverteilung im Schnee, Eis und Wasser zustande und inwieweit ist daran die Strahlungsdurchlässigkeit beteiligt? Im allgemeinen wurde in früheren Untersuchungen der Bedeutung der Strahlungsdurchlässigkeit für das Zustandekommen bestimmter Temperaturverteilungen nur geringe Beachtung geschenkt. So verwendete man z. B. zur Berechnung der Wärmeleitfähigkeit im Schnee mit Hilfe der beobachteten Temperaturänderungen in den verschiedenen Tiefen die klassische Formel für die Verlagerung solcher Temperaturwellen von der Oberfläche zur Tiefe hin (siehe z. B. KERÄNEN 1929), und auch bei Wasser wurde die Temperaturwelle als von der Oberfläche ausgehend betrachtet (SCHMIDT, W., 1925). Bei allen diesen Problemen des Wärmehaushaltes müssen jedoch Theorie und Beobachtung Hand in Hand gehen, da in die theoretischen Formeln eine Reihe von Grössen eingehen, die durch direkte Messung bestimmt werden müssen. Die Beobachtungen lieferten nun gerade in den strahlungsdurchlässigen Medien wie Schnee und Wasser gewisse besondere Merkmale, die durch die klassische Theorie keine Erklärung fanden. So zeigte sich, dass während des Einstrahlungsprozesses im Schnee das

Maximum der Temperatur nicht an der Oberfläche sondern einige cm tiefer auftritt (KERÄNEN 1929) und im Wasser wurde das Phänomen der sogenannten Sprungschicht entdeckt. Während bereits KERÄNEN vermutete, dass die Tiefenverlagerung des Temperaturmaximums bei Schnee eine Folge seiner Strahlungsdurchlässigkeit sei, glaubte man bei Wasser eine Erklärung für die Sprungschicht durch die von der Oberfläche zur Tiefe hin rasch abnehmende turbulente Durchmischung geben zur können. (MUNK 1948). Will man eine Theorie des Wärmehaushaltes strahlungsdurchlässiger Medien aufstellen, so ist es zweckmässig zunächst von einfachen Verhältnissen auszugehen. Solche relativ einfache Bedingungen liegen bei Schnee vor. Schnee kann in erster Näherung als für die Sonnenstrahlung nicht selektiv absorbierendes Medium angesehen werden, so dass man also in diesem Fall mit einem einheitlichen (endlichen) Absorptionskoeffizienten für die Einstrahlung rechnen kann. Andrerseits ist Schnee, wie auch Wasser und Eis für die langwellige Gegenstrahlung der Atmosphäre praktisch undurchlässig, das heisst diese Medien können für die in der Natur vorkommende Temperaturstrahlung als nahezu schwarzer Körper betrachtet werden. Diese Tatsache, dass für die in der Natur vorkommenden Temperaturen Schnee, Eis und Wasser für den Einstrahlungsvorgang soweit es sich dabei um die Sonnen- und Himmelsstrahlung handelt durchlässig, jedoch für die ausgesandte Strahlung undurchlässig sind, ist das wesentliche Merkmal der ganzen Theorie. Ich habe an anderer Stelle (REUTER 1948 a) mit Hilfe dieses Ansatzes eine Theorie des Wärmehaushaltes von Schnee entwickelt, die eine Reihe von beobachteten Merkmalen der Temperaturverteilung zu erklären gestattet. In der folgenden Untersuchung soll nun ein Ansatz für eine allgemeine Theorie selektiv strahlungsdurchlässiger Medien durchgeführt werden mit besonderer Anwendung dieser Theorie auf die Verhältnisse in Eis, das ein selektiv absorbierendes Medium ähnlich dem Wasser darstellt, bei dem aber die Komplizierung durch die turbulente Durchmischung wegfällt. Immerhin lassen sich aber die so für Eis abgeleiteten Ergebnisse auch in erster Näherung auf ruhiges Wasser anwenden. Da Temperaturbeobachtungen in Eis m. W. keine vorliegen, so kam es mir vor allem darauf an, die durch die Theorie aufgezeigten Unterschiede im Erwärmungsprozess verschiedener Medien (Boden, Schnee, Eis) herauszuarbeiten, so dass die gewonnenen Ergebnisse zunächst mehr qualitativ zu werten sind.

Erwärmung eines selektiv strahlungsdurchlässigen Mediums durch konstante kurzwellige Einstrahlung bei gleichzeitiger konstanter effektiver Ausstrahlung

Für unsere theoretischen Überlegungen setzen wir im nachfolgenden voraus: Eine gewisse Zeit t lang sei eine konstante Sonneneinstrahlung wirksam. Für die langwellige Eigenstrahlung der Atmosphäre möge das bestrahlte Medium praktisch undurchlässig sein. Ebenso weise das Medium in demjenigen Wellenlängenbereich, der der ausgesandten langwelligen Temperaturstrahlung vornehmlich zukommt, eine unendlich grosse Absorption auf. Dann kann die »effektive« Strahlungstemperatur des Mediums gleich der Oberflächentemperatur gesetzt werden und die tatsächlich wirksame Ausstrahlung der Oberfläche nach Abzug der langwelligen Gegenstrahlung der Atmosphäre in Form der sogenannten »effektiven« Ausstrahlung R angesetzt werden. Zur Vereinfachung nehmen wir an, dass diese Grösse R einen nahezu konstanten Betrag in dem betreffenden Zeitraum aufweist, wodurch die Rechnung eine wesentliche Verkürzung erfährt. Über die Berechtigung einer solchen Voraussetzung will ich hier nicht weiter diskutieren, da ich dies bereits in der erwähnten Arbeit getan habe. Jedenfalls ist durch Messungen belegt, dass die effektive Ausstrahlung in kürzeren Zeiträumen nur geringe Schwankungen zeigt (siehe dazu REUTER 1947 u. 1948 a, HANN-SÜRING 1937 mit den dort vorhandenen Literaturhinweisen). Die nachfolgenden Überlegungen behalten ihre Gültigkeit, wenn wir unter R nicht allein die effektive Ausstrahlung sondern den gesamten Wärmeverlust der Oberfläche durch Strahlung, Wärmeleitung, Austausch zur Luft hin und Verdunstung verstehen, sofern nur die Konstanz dieses Betrages gewährleistet ist. Durch Miteinbeziehung dieser Wärmeprozesse in die Oberflächenbilanz würde aber nur eine angenäherte Berücksichtigung dieser wesentlich vom Temperaturgradienten Erde-

Luft abhängigen Wärmetransporte stattfinden, so dass wir davon absehen wollen. Dies ist umso mehr berechtigt, als wir unsere Wärmehaushaltsbetrachtungen in Eis und Schnee bei tiefen Temperaturen durchführen, wo die Verdunstung praktisch vernachlässigbar ist und wir nur als weitere Einschränkung ruhende Luft über dem Medium voraussetzen müssen. Wie eine exakte Behandlung derartiger Probleme bei Mitberücksichtigung des Austausches in der Luft durchgeführt werden muss, habe ich an anderer Stelle für den nächtlichen Abkühlungsprozess gezeigt (REUTER 1947).

Für die Rechnung zerlegen wir die einstrahlende kurzwellige Sonnenstrahlung in $n+1$ Spektralbereiche, die hinreichend klein gewählt seien, so dass innerhalb eines jeden Bereiches mit einem einheitlichen Absorptionskoeffizienten gerechnet werden kann. Mithin erhalten wir für die die horizontale Oberfläche erreichende Strahlung I:

$$I = i_0 + i_1 + \ldots\ldots i_n = \sum_0^n i_\lambda \qquad (1)$$

Um allgemein zu bleiben, lassen wir die Möglichkeit einer selektiven Reflexion an der Oberfläche zu und bezeichnen die Reflexzahlen mit $\alpha_0, \alpha_1 \ldots \alpha_n$. Setzen wir entsprechend $1 - \alpha_\lambda = \beta_\lambda$, und sind die den einzelnen Spektralbereichen zukommenden Absorptionskoeffizienten $a_0,\ a_1 \ldots a_n$, so wird im Niveau z (positiv von der Oberfläche zur Tiefe hin gerechnet) in der Zeiteinheit eine durch die eindringende Strahlung hervorgerufene Temperaturänderung von:

$$\frac{\partial T}{\partial t} = \frac{1}{\varrho c} \sum_0^n a_\lambda \beta_\lambda i_\lambda e^{-a_\lambda z} \qquad (2)$$

stattfinden, wenn ϱ und c die Dichte bezw. die spezifische Wärme des Mediums darstellen und T die Temperatur bedeutet.

Zusammen mit der (eindimensionalen) Wärmeleitungsgleichung ergibt dies, wenn wir unter $K = \dfrac{k}{\varrho c}$ die Temperaturleitfähigkeit (k = Wärmeleitfähigkeit) verstehen folgende Differentialgleichung:

$$\frac{\partial T}{\partial t} = K \frac{\partial^2 T}{\partial z^2} + \frac{1}{\varrho c} \sum_0^n a_\lambda \beta_\lambda i_\lambda e^{-a_\lambda z} \qquad (3)$$

wobei wir als Nebenbedingungen gemäss den früheren Ausführungen:

$$t = 0,\ z > 0;\ T = T_a$$

$$t > 0,\ z = 0;\ k \left(\frac{\partial T}{\partial z}\right)_{z=0} = R = \text{const} \qquad (3a)$$

setzen, mithin eine Isothermie vom Betrage T_a als Anfangszustand annehmen.

Zur Lösung von (3) mit (3 a) zerlegen wir die Differentialgleichung entsprechend den $n+1$ Spektralbereichen in $n+1$ Teilgleichungen, was infolge der Linearität möglich ist und erhalten für die erste derartige Gleichung:

$$\frac{\partial T_0}{\partial t} = K \frac{\partial^2 T_0}{\partial z^2} + \frac{a_0 \beta_0 i_0}{\varrho c} e^{-a_0 z} \qquad (4)$$

mit den Nebenbedingungen:

$$t = 0,\ z > 0;\ T_0 = T_a$$

$$t > 0,\ z = 0;\ k \left(\frac{\partial T_0}{\partial z}\right)_{z=0} = R = \text{const} \qquad (4a)$$

während für die übrigen n Gleichungen:

$$\frac{\partial T_\lambda}{\partial t} = K \frac{\partial^2 T_\lambda}{\partial z^2} + \frac{a_\lambda \beta_\lambda i_\lambda}{\varrho c} e^{-a_\lambda z} \qquad (5)$$

$$\lambda = 1,\ 2,\ \ldots\ n$$

mit den Nebenbedingungen:

$$t = 0,\ z > 0;\ T_\lambda = 0$$

$$t > 0,\ z = 0;\ \left(\frac{\partial T_\lambda}{\partial z}\right)_{z=0} = 0 \qquad (5a)$$

gesetzt werden muss.

Die Lösung von (4) mit (4 a) erfolgt (siehe REUTER 1948 a) durch den Ansatz:

$$T_0 = \gamma - q\ e^{-a_0 z} \text{ mit } q = \frac{\beta_0 i_0}{k\, a_0}, \qquad (6)$$

so dass dann in der Variablen γ folgende Differentialgleichung resultiert:

$$\frac{\partial \gamma}{\partial t} = K \frac{\partial^2 \gamma}{\partial z^2} \qquad (7)$$

mit

$$t = 0,\ z > 0;\ \gamma = q\, e^{-a_0 z}$$

$$t > 0,\ z = 0;\ \left(\frac{\partial \gamma}{\partial z}\right)_{z=0} = \frac{R - \beta_0 i_0}{k} \qquad (7a)$$

Wie hier nicht näher ausgeführt werden soll, lässt sich die endgültige Lösung von (4) mit (4 a) schliesslich in der Form:

$$\Delta T_0 = T_0 - T_a = \frac{\beta_0 i_0}{k a_0} F(a_0, z, t) + \\ + (\beta_0 i_0 - R)\, G(z, t) \qquad (8)$$

darstellen. Hierin

$$2 F(a_0, z, t) =$$

$$= \left\{ e^{K a_0^2 t - a_0 z}\left[1 - \Theta\left(a_0\sqrt{Kt} - \frac{z}{2\sqrt{Kt}}\right)\right] + \right.$$

$$\left. + e^{K a_0^2 t + a_0 z}\left[1 - \Theta\left(a_0\sqrt{Kt} + \frac{z}{2\sqrt{Kt}}\right)\right] - e^{-a_0 z}\right\} \qquad (9)$$

und

$$G(z, t) = \frac{2\sqrt{t}}{\sqrt{\pi k \varrho c}}\, e^{-\frac{z^2}{4Kt}} - \frac{z}{k}\left[1 - \Theta\left(\frac{z}{2\sqrt{Kt}}\right)\right] \qquad (10)$$

wenn wir unter

$$\Theta(x) \equiv \frac{2}{\sqrt{\pi}} \int_0^x e^{-\zeta^2}\, d\zeta \qquad (11)$$

das bekannte Gaussche Fehlerintegral verstehen.

Dass (8) mit (9) und (10) tatsächlich die gesuchte Lösung von (4) mit allen notwendigen Nebenbedingungen (4 a) ist, kann nachträglich leicht durch differenzieren bewiesen werden. Für (5) mit (5 a) ergibt sich analog die Lösung, wenn wir in (8) $R = 0$ setzen, zu:

$$\Delta T_\lambda = \frac{\beta_\lambda i_\lambda}{k a_\lambda} F(a_\lambda, z, t) + \beta_\lambda i_\lambda G(z, t), \qquad (12)$$

so dass endlich die Gesamtlösung von (3) mit (3 a) unter Berücksichtigung von (1) in der Form:

$$\Delta T = \frac{1}{k} \sum_0^n \frac{\beta_\lambda i_\lambda}{a_\lambda} F(a_\lambda, z, t) + (\beta I - R)\, G(z, t) \qquad (13)$$

geschrieben werden kann, wenn hier $1 - \beta$ die Reflexzahl für die Gesamtintensität I bedeutet. Für $a_\lambda \to \infty$ verschwindet der erste Term in (13). Er stellt mithin den Anteil der Temperaturänderung dar, der durch die Strahlungsdurchlässigkeit des Mediums hervorgerufen wird. In einem strahlungsundurchlässigen Medium z. B. im Erdboden tritt nur der zweite Term in Erscheinung.

Näherung der Funktion $F(a_\lambda, z, t)$ für Sonderfälle

Ist der Ausdruck $a_\lambda\sqrt{Kt}$ nach hinreichend langer Zeit genügend gross, was vor allem vom Absorptionskoeffizienten a_λ abhängen wird, können auch die beiden Ausdrücke

$$a_\lambda\sqrt{Kt} - \frac{z}{2\sqrt{Kt}} \text{ bezw. } a_\lambda\sqrt{Kt} + \frac{z}{2\sqrt{Kt}} \text{ als so}$$

gross vorausgesetzt werden, dass eine Näherung des Fehlerintegrals mit Hilfe einer Reihenentwicklung möglich ist. In der bereits mehrfach erwähnten Arbeit ist dies für den Fall eines einheitlichen Absorptionskoeffizienten durchgeführt worden, so dass wir uns hier nur darauf beschränken, die Verallgemeinerung auf eine selektive Absorption durchzuführen. Unter der Annahme, dass bereits das erste Glied der Reihenentwicklung des Fehlerintegrals eine hinreichende Näherung darstellt, wird (siehe RIEMANN-WEBER 1901 Bd. I. § 26):

$$F(a_\lambda, z, t) = \frac{1}{\sqrt{\pi}}\, e^{-\frac{z^2}{4Kt}} \cdot \frac{1}{a_\lambda\sqrt{Kt}} - e^{-a_\lambda z} \qquad (14)$$

Lassen wir die Zeit noch weiter ansteigen, so wird in (14) auch der erste Term gegenüber dem zweiten vernachlässigbar und wir erhalten:

$$F(a_\lambda, z, t) = -\, e^{-a_\lambda z} \qquad (14a)$$

Diese Form der Funktion $F(a_\lambda, z, t)$ kann verwendet werden, um die Bedingungsgleichung für diejenige Tiefe z_0 aufzustellen, in der nach hinreichend langer Zeit (stationärer Zustand) das Temperaturmaximum auftritt.

Man hat dazu nur in (13) die vereinfachte Form (14 a) einzusetzen und nach z zu differenzieren. Das so erhaltene Resultat gleich Null gesetzt, ergibt die gesuchte Bedingungsgleichung. Wie man nach leichten Umformungen unter Berücksichtigung der zu (14 a) führenden Vereinfachungen bezw. Vernachlässigungen berechnet, lautet diese Bedingungsgleichung:

$$\sum_{0}^{n} \beta_\lambda i_\lambda e^{-a_\lambda z_0} = \beta I - R = B \qquad (15)$$

wenn wir zur Abkürzung mit B die Strahlungsbilanz der Oberfläche bezeichnen.

Für den Betrag des Temperaturmaximums, das heisst für den Temperaturunterschied zwischen Oberfläche und der Tiefe z_0, ergibt sich durch Einsetzen von (15) in (13):

$$T(z_0) - T(0) = \frac{1}{k}\left[\sum_{0}^{n} \frac{\beta_\lambda i_\lambda}{a_\lambda} - \sum_{0}^{n} \frac{\beta_\lambda i_\lambda}{a_\lambda} e^{-a_\lambda z_0}\right] - \frac{z_0 B}{k} \qquad (16)$$

Für den Sonderfall $a_0 = a_1 = a_2 = \ldots a_n = a$ ergibt sich aus (15) und (16) das Resultat für ein nicht selektiv absorbierendes Medium, nämlich:

$$\beta I e^{-a z_0} = B \text{ oder } z_0 = \frac{1}{a} \ln \frac{\beta I}{B} \qquad (15a)$$

und

$$T(z_0) - T(0) = \frac{1}{ka}\left[R - B \cdot \ln \frac{\beta I}{B}\right] \qquad (16a)$$

Dieser Fall findet bei Schnee Anwendung und wurde von mir in der bereits erwähnten Arbeit eingehend diskutiert. Eine Anwendung der allgemeinen Formeln (15) und (16) auf ein spezielles Problem, nämlich den Wärmehaushalt der menschlichen Haut, wurde von mir an anderer Stelle (REUTER 1948 b) gegeben.

Eine andere Näherung für die Funktion $F(a_\lambda, z, t)$ ist möglich, wenn der Absorptionskoeffizient a_λ sehr klein ist und gleichzeitig der Ausdruck $\sqrt{Kt}$ nicht zu gross wird. Zum Unterschied von der ersten Näherung werden also hier Erwärmungsprozesse von relativ kurzer Zeitdauer in Betracht kommen bei grosser Strahlungsdurchlässigkeit des Mediums. Für hinreichend grosse Werte von z nämlich nimmt die Funktion $\Theta\left(a_\lambda \sqrt{Kt} - \frac{z}{2\sqrt{Kt}}\right)$ genügend genau den Wert minus eins an, während gleichzeitig die Funktion $\Theta\left(a_\lambda \sqrt{Kt} + \frac{z}{2\sqrt{Kt}}\right)$ dem Betrag von plus eins sehr nahe kommen wird. Berücksichtigen wir dies in (9), so erhalten wir:

$$F(a_\lambda, z, t) = e^{-a_\lambda z}\left(e^{K a_\lambda^2 t} - 1\right) \qquad (17)$$

Da unter der Annahme $a_\lambda \ll 1$ für die e-Potenz in der Klammer $1 + K a^2 t$ geschrieben werden kann, ergibt sich als weitere Vereinfachung:

$$F(a_\lambda, z, t) = a_\lambda^2 K t \cdot e^{-a_\lambda z} \qquad (17a)$$

Diese Näherung findet Anwendung, wenn wir bei grosser Strahlungsdurchlässigkeit, also etwa bei Eis und Wasser im kurzwelligen Teil des Spektrums, nach nicht zu langer Zeit die Temperaturänderung (Erwärmung) in grösserer Tiefe berechnen wollen. Nach (13) ergibt sich nämlich in diesem Fall, da die Funktion $G(z, t)$ für genügend grosse Werte von z gegen Null geht, unter Verwendung von (17 a):

$$\Delta T = \sum_{0}^{n} \frac{\beta_\lambda i_\lambda a_\lambda}{\varrho c} \cdot t \cdot e^{-a_\lambda z} \qquad (18)$$

also bei vorgegebener Einstrahlung eine mit der Zeit direkt proportionale jedoch von der Ausstrahlung der Oberfläche unabhängige Temperaturerhöhung.

Anwendung der theoretischen Ergebnisse

Wir wollen nun die im Vorstehenden abgeleiteten theoretischen Ergebnisse auf praktische in der Natur realisierte Vorgänge anwenden. Wir nehmen drei verschiedene Medien an und zwar:

1. Feuchter Sand: $a_0 = a_1 = \ldots a_n = \infty$
   Albedo: 10 %, $\sqrt{k\varrho c} = 3{,}30 \cdot 10^{-2}$ im CGS-System, $k = 0{,}0032$ cal/cm sec grad (siehe HANN-SÜRING 1937)

2. Altschnee: $a_0 = a_1 = \ldots\ldots a_n = a = 0{,}15\ \mathrm{cm}^{-1}$

   Albedo: 60 %, $\varrho = 0{,}4\ \mathrm{g\ cm}^{-3}$,
   $k = 0{,}00109$ cal/cm sec grad
   $c = 0{,}505$ cal/g grad

   Wir nehmen Schnee grösserer Dichte, um wenigstens einigermassen mit einer einheitlichen Wärmeleitung rechnen zu können (siehe REUTER 1948 a).

3. Eis: Als selektiv strahlungsdurchlässiges Medium handelt es sich hier, für die einzelnen Wellenlängenbereiche die entsprechenden Absorptionskonstanten festzulegen. Wir setzen voraus, dass bei Klareis dieselben Verhältnisse wie bei Wasser vorliegen (siehe dazu BODE 1909). In der Tabelle I sind die entsprechenden Werte nach EMDEN (1940) und ASCHKINASS (1895) zusammengestellt.
   Albedo: keine. Dies erscheint begründet, wenn wir senkrechte Einstrahlung voraussetzen, was wir tun wollen. Überdies handelt es sich hier vor allem um einen relativen Vergleich zu den beiden anderen Medien, bei denen dann also die Albedowerte relativ zu Klareis (bei senkrechter Bestrahlung) zu verstehen wären. $\varrho = 0{,}917\ \mathrm{g\ cm}^{-3}$, $k = 0{,}0051$ cal/cm sek grad $c = 0{,}505$ cal/g grad.

Für alle drei Fälle setzen wir voraus, dass durch drei Stunden eine Gesamteinstrahlung gemäss Tabelle I von $I = 0{,}79\ \mathrm{cal/cm}^2$ Min wirksam sei, während gleichzeitig der Verlust der Oberfläche durch effektive Ausstrahlung $R = 0{,}10\ \mathrm{cal/cm}^2$ Min betrage. Als Anfangszustand ist in allen drei Fällen dieselbe Isothermie angenommen.

Das Ergebnis ist in den Abbildungen 1 bis 3 veranschaulicht. Wie wir in den theoretischen Betrachtungen gesehen haben, gibt die Funktion $F(a_i, z, t)$ in (13) den Anteil in der gesamten Temperaturänderung, der gerade durch die Strahlungsdurchlässigkeit des Mediums bedingt ist, wieder. Beim Erdboden ist die Funktion Null für alle Zeiten. Um nun in den Abbildungen die Bedeutung der Strahlungsdurchlässigkeit für den Wärmehaushalt besonders augenscheinlich zu machen, ist in Abb. 2 und 3 also für Schnee und Eis die Funktion $(\beta I - R)$. $G(z, t)$ strichliert nach (13) und (10) eingezeichnet. Diese Kurven stellen also diejenige Temperaturänderung dar,

Tabelle 1

Absorptionskoeffizienten von Wasser (Eis) für verschiedene Wellenlängenbereiche mit den entsprechenden Strahlungsintensitäten an der Erdoberfläche bei Annahme einer Einstrahlung von 1 cal/cm² Min an der Grenze der Atmosphäre auf die horizontale Fläche (nach EMDEN).

$\Delta\lambda$ μ	a_λ cm^{-1}	j_λ cal/cm² Min	p_λ	i_λ cal/cm² Min
0,00—0,45	0,0002	0,120	0,530	0,0636
0,45—0,50	0,0002	0,087	0,674	0,0586
0,50—0,55	0,00025	0,079	0,719	0,0566
0,55—0,60	0,0006	0,073	0,746	0,0544
0,60—0,65	0,0021	0,065	0,781	0,0508
0,65—0,70	0,0036	0,056	0,820	0,0459
0,70—0,75	0,015	0,049	0,848	0,0415
0,75—0,80	0,023	0,044	0,860	0,0377
0,80—0,85	0,032	0,038	0,870	0,0331
0,85—0,90	0,052	0,033	0,888	0,0284
0,90—1,00	0,24	0,055	0,893	0,0491
1,00—1,15	0,28	0,062	0,906	0,0561
1,15—1,35	1,17	0,051	0,915	0,0540
1,35—1,85	23,0	0,091	0,902	0,0820
1,85—2,5	67,0	0,033	0,900	0,0297
2,5 —3,5	10^3	0,057	0,900	0,0513
3,5 — ∞	$5 \cdot 10^2$			
	Summe	1,000		0,7928

$\Delta\lambda$ = Wellenlängenbereich, j_λ = Strahlungsintensität an der Grenze der Atmosphäre, p_λ = Transmissionskoeffizient der Atmosphäre, i_λ = Strahlungsintensität an der Erdoberfläche.

die ein Medium unter den gegebenen Bedingungen erfahren würde, das zwar dieselben Wärmeleitungs- und Strahlungseigenschaften wie Schnee bezw. Eis aufweist, für die Strahlung jedoch undurchlässig ist. Man kann an Hand der Abbildungen das Rechenergebnis in folgender Weise überprüfen: Wir bestimmen den Wärmeinhalt, der dem betreffenden Medium durch die Einstrahlung zugeführt wird. Dies geschieht derart, dass wir in Abb. 1 bis 3 den Flächeninhalt zwischen der Ordinate Null und der Tautochrone ausmessen und mit der (mittleren) Dichte und spezifischen Wärme multiplizieren. Verwenden wir in Abb. 2 und 3 die strichlierten Kurven, so erhalten wir auf diese Weise die drei Grössen W_1, W_2 und W_3. Bezeichnen wir die den drei Fällen entsprechenden Strahlungsbilanzen der Oberfläche bei Berücksichtigung der verschieden vorausgesetzten Albedo mit resp. B_1, B_2 und B_3, so muss

$$W_1 : W_2 : W_3 = B_1 : B_2 : B_3 \qquad (19)$$

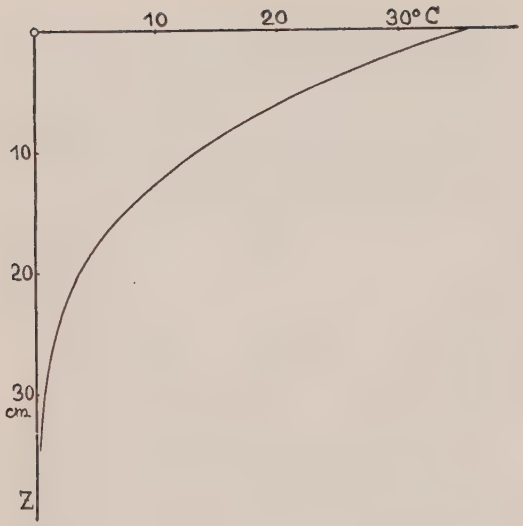


Abbildung 1. Erwärmung von Erdboden (feuchter Sand) als Funktion der Tiefe z nach einer dreistündigen Einstrahlung vom Betrage $I_0 = 0{,}79$ cal/cm² Min bei gleichzeitiger effektiver Ausstrahlung von $R = 0{,}1$ cal/cm² Min unter der Annahme einer Albedo von 10 %.

sein. Eine derartige Überprüfung liefert innerhalb der Fehlergrenzen der Planimetrierung tatsächlich die Beziehung (19). Mit einer ähnlichen Methode kann man aber auch die

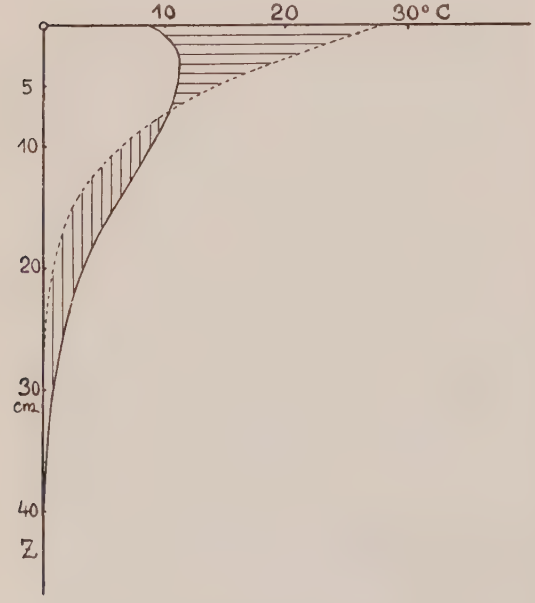


Abbildung 2. Erwärmung von Altschnee mit einer Dichte von 0,4 g/cm³ und einer Albedo von 60 % unter denselben Bedingungen wie bei Abb. 1.

Wirkung der Strahlungsdurchlässigkeit im Fall 2 und 3, die dann zu den ausgezogenen Tautochronen in Abb. 2 und 3 führt, feststellen und die Theorie nachprüfen. So muss offenbar in Abb. 2 eine Flächengleichheit der schraffierten Flächen vorliegen, was augenscheinlich zeigt, wie durch die Strahlungsdurchlässigkeit eine Tiefenverlagerung des Wärmeinhaltes auftritt. Das lässt sich durch Planimetrierung leicht nachweisen. Dieselben Verhältnisse finden wir auch im Fall 3 vor, doch erscheint im ersten Moment hier die

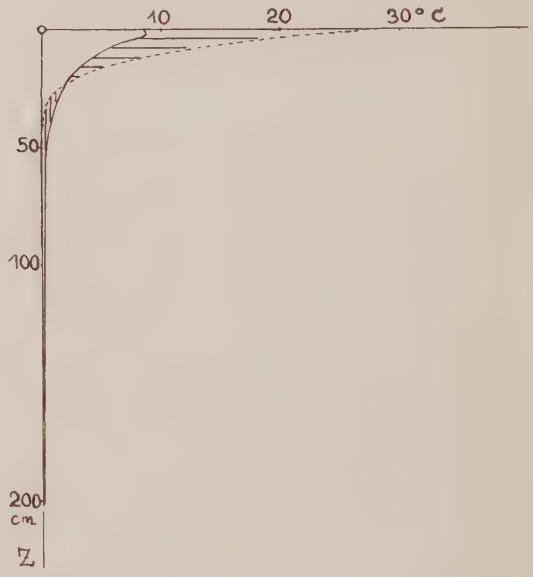


Abbildung 3. Erwärmung von Eis mit einer selektiven Absorption entsprechend den Werten der Tabelle I unter denselben Bedingungen wie in Abb. 1 und 2 ohne Albedo.

Theorie ein unrichtiges Ergebnis zu liefern. Es ist aber zu bedenken, dass bei Eis durch die ausserordentlich geringen Absorptionskoeffizienten im kurzwelligen Teil des Spektrums (siehe Tabelle I) ein beachtlicher Teilbetrag der Gesamtstrahlung bis in grosse Tiefen eindringt, ohne wesentlich geschwächt zu werden, und dort zu einer, wenn auch nur geringen, Erwärmung führt, die in dem Masstab der Abb. 3 nicht ersichtlich wird. Um diese Verhältnisse genauer studieren zu können, müssen wir auf die Näherungsformel (18) zurückgreifen. Danach berechnet sich für Eis noch in einer Tiefe von 10 Metern eine Erwärmung von 0,026 Grad und in 15

Meter eine solche von 0,018 Grad. Infolge der grossen Vertikalausdehnung dieser an sich geringfügigen Erwärmung resultiert für den Gesamtwärmeinhalt dennoch ein fühlbarer Beitrag, der so viel ausmacht, dass die Planimetrierung analog dem Fall 2 die Gleichheit der entsprechenden Flächen liefert.

Um aber die Verhältnisse im Fall 3 klarer erkennen zu können, wollen wir das Problem noch von einer anderen Seite her betrachten. Wir fragen uns, in welcher Tiefe beim Eis dieselbe Erwärmung auftritt, die ein sonst

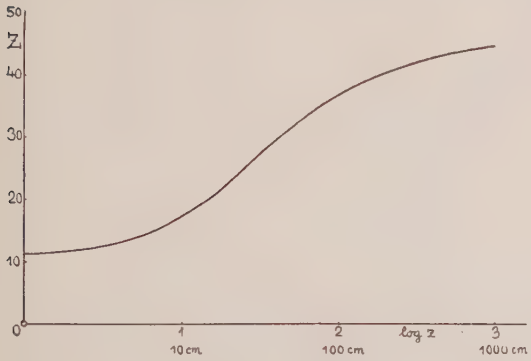


Abbildung 4. Kurve gleicher Erwärmung von strahlungsdurchlässigem und strahlungsundurchlässigem Eis unter denselben Bedingungen wie in Abb. 3.

gleichbeschaffenes aber »strahlungsundurchlässiges» Medium aufweisen würde. Das Ergebnis ist in Abbildung 4 dargestellt. Dabei wurde auf der Abszisse eine logarithmische Einteilung gewählt, um den Masstab zu verkleinern. Man erkennt aus dieser Abbildung sehr eindrucksvoll, welche Wirkung die Strahlungsdurchlässigkeit auf die Tiefenverlagerung der Erwärmung ausübt. Eine Erwärmung, die sonst in 45 cm Tiefe auftreten würde, findet sich in Wirklichkeit noch in 10 m.

Wir haben hier nur einen ganz speziellen Fall des Erwärmungsprozesses unter Annahme wesentlicher Vereinfachungen betrachtet und es drängen sich sofort eine ganze Reihe von Fragen über eine mögliche Erweiterung der obigen Theorie auf. So wäre zunächst der Fall einer periodischen Einstrahlung zu untersuchen, es wäre die Frage der Wärmeabgabe der Oberfläche des Mediums an die Luft durch Austauschvorgänge in die Theorie miteinzubeziehen, es wäre zu überlegen, inwieweit die Annahme einer veränderlichen effektiven Ausstrahlung auf das Ergebnis modifizierend wirkt und dergl. mehr. Für den Fall einer nichtselektiven Absorption habe ich einige dieser Fragen bereits in der mehrfach erwähnten Arbeit (REUTER 1948 a) behandelt. Eine Verallgemeinerung auf ein Medium mit selektiver Absorption ist jedoch im allgemeinen keineswegs so leicht durchzuführen, wie dies hier möglich war, da die Formeln eine zu komplizierte Gestalt annehmen. Eine qualitative Abschätzung, die vor allem die Unterschiede der strahlungsdurchlässigen Medien gegenüber den strahlungsundurchlässigen aufzeigt, wird aber bereits durch die hier durchgeführte einfache Theorie gegeben.

Eine andere Erweiterung der durchgeführten Rechnungen, nämlich der Übergang zum Studium des Wärmehaushaltes flüssiger strahlungsdurchlässiger Medien, insbesondere dem Wärmehaushalt von Wasser, stösst auf gewisse Schwierigkeiten, da bei Wasser neben der molekularen Wärmeleitung noch diejenige durch Austausch und Konvektion eine überragende Rolle spielt. Beschränken wir uns allerdings auf absolut ruhiges Wasser und betrachten wir einen reinen Erwärmungsprozess im Sinne der oben angeführten Theorie, so dass eine Konvektion nicht in Frage kommt, dann können wir die Ergebnisse, die für Eis abgeleitet wurden, ohne weiteres auch auf Wasser anwenden. Die gegenüber Eis geänderte Dichte und spezifische Wärme führt zu keiner grundsätzlichen Änderung des Ergebnisses der Abb. 3. Insbesondere ist die Näherungsformel (18) auch für grosse Wassertiefen verwendbar. Allerdings muss beachtet werden, dass die Werte der Tabelle I für chemisch reines Wasser gelten und in der Natur nur sehr begrenzte Gültigkeit aufweisen, wie hier durch Hinweis auf SAUBERER und RUTTNER (1941) betont sei.

Abschliessend müssen wir aber auf eine besondere durch die hier durchgeführten Rechnungen und aus Abbildung 2 und 3 ersichtliche Erscheinung hinweisen, zumal diesbezüglich vielfach in der Literatur unrichtige Vorstellungen zu finden sind. Zunächst lassen die Abbildungen 2 und 3 das formelmässig in (15) und (16) bezw. (15 a) und (16 a) wiedergegebene Temperaturmaximum in einigen cm Tiefe erkennen, das nach dem eben gesagten ähnlich wie bei Eis auch bei Wasser in etwa 1 cm Tiefe zu finden wäre. Das heisst die Strahlungsdurchlässigkeit bewirkt knapp unter

der Oberfläche eine grössere Erwärmung als an der Oberfläche selbst. Wenn wir demgegenüber bei W. SCHMIDT (1925) lesen, dass die Temperaturerhöhung aus Absorption im reinen Wasser schon in ein Zentimeter Tiefe gegenüber der Erwärmung an der Oberfläche zu vernachlässigen ist und dass die von der Sonnenstrahlung hervorgerufene Temperaturwelle als von der Oberfläche ausgehend betrachtet werden kann, so finden wir hier einen Widerspruch mit der von uns abgeleiteten Theorie. Ohne jedoch näher auf die Verhältnisse im Wasser, die einer eigenen Abhandlung vorbehalten bleiben sollen, einzugehen, sei nur abschliessend dazu betont, dass mithin auch die von W. SCHMIDT auf Grund seiner Vorstellung für Wasser abgeleiteten thermischen Austauschkoeffizienten bei weitem zu gross sind, wenn wir, wie dies hier geschehen ist, die Strahlungsdurchlässigkeit als wesentlichen Faktor für die Erwärmung tieferer Schichten ansehen.

LITERATUR

ASCHKINASS, E. 1895: Untersuchungen des Absorptionsspektrums von Wasser. *Wied. Ann.* **55**, S. 401.

BODE, G. 1909: Verhalten des Eises im ultraroten Spektrum. *Ann. Phys.* 4 R Bd. 30.

EMDEN, R. 1940: Zum Temperaturproblem der Seen. *Helv. Phys. Acta* **8**, Fasc. V, S. 396.

HANN-SÜRING 1937: Lehrbuch der Meteorologie 5. Aufl. S. 104 ff. Leipzig.

KERÄNEN, J. 1929: Wärme- und Temperaturverhältnisse der obersten Bodenschichten. Naturw. Monographien u. Lehrb. 8, II. Teil, S. 169 ff, Berlin.

MUNK, W. H. and ANDERSON, E. R. 1948: »Notes on a theory of thermocline. *J. Mar. Res.* **7** (3), S. 276.

REUTER, H. 1947: Zur Theorie der nächtlichen Abkühlung der bodennahen Schicht und Ausbildung der Bodeninversion. *Sitz.Ber. Akad. Wiss.* Wien, Abt. II a 155.

REUTER, H. 1948 a: Ueber die Theorie des Wärmehaushaltes einer Schneedecke. *Arch. Met. Geophys. u. Bioklim.* **A 1**, S. 62.

—— 1948 b: Ueber die Tiefenverlagerung des Temperaturmaximums bei Bestrahlung der menschlichen Haut. Wetter u. Leben *Zeitschr. f. prakt. Bioklim.* **1**, H. 9.

RIEMANN-WEBER 1901: Die partiellen Differentialgleichungen der Physik. Braunschweig 2. Aufl.

SAUBERER-RUTTNER 1941: Die Strahlungsverhältnisse der Binnengewässer. Probl. d. kosm. Physik, Akad. Verl. Ges. Leipzig.

SCHMIDT, W. 1925: Der Massenaustausch in freier Luft und verwandte Erscheinungen. Probl. d. kosm. Physik, Hamburg.

The Problem of Artificial Control of Rainfall on the Globe[1]

II. The Coastal Orographic Maxima of Precipitation in Autumn and Winter

By TOR BERGERON, University of Uppsala

(Manuscript received 16 April 1949.)

Abstract

The coastal maxima of rainfall during autumn and winter cannot be explained alone by the classical theory of orographic precipitation. Three different *frictional effects*, and one effect leading to *orographically conditioned convergence* at quasi-isophysic flow, are discussed and illustrated by Scandinavian and Dutch rainfall maps of remarkable special cases of maximum coastal precipitation. The rainfall maps over Holland even hint secondary maxima further inland, at zero-level and parallel to the coastal maximum; they might be explained in connection with standing waves in the sense of QUENEY and others. With all these effects a *low condensation level* and an *efficient precipitation release*, by ice-nuclei or in other ways, seem indispensable to cause intense orographic condensation and bring its results promptly down to the earth.

As stated in article I such orographic cloud-systems, if lacking an efficient release, just represent the clouds where artificial seeding might materially increase precipitation, because (1) condensation may for a long time keep intense and stationary, and (2) the clouds may keep releasable but unreleased. This increase, if at all possible, would diminish precipitation inland where it is much more useful. — A better scheme would therefore be trying to overseed coastal cloud-systems in order to prevent their precipitation release, and thereby increasing precipitation inland, provided the overseeding could be avoided there. This overseeding process might also be used for dissipating inland Cunb, but presumably not for dissipating tropical hurricanes.

The generally accepted idea of the *increase of precipitation with height* is often, e.g. in Norway, contradicted by individual cases with a marked decrease from the very coast and upward when moving inland. A rational attack on this problem was hitherto impossible because the complex nature of orographic precipitation (hinted above) was not recognized. — Outside the Tropics *excessive precipitation* needs low cloud formed by an intense updraft at high temperature (= maximum condensation and minimum evaporation) in continual contact with an upper cloud (part) surpassing the — 10°C-level, or seeded artificially (= efficient precipitation release). These conditions will be fulfilled in certain orographic clouds discussed above, and within some convective and frontal clouds.

---

1 Based partly on a lecture delivered Aug. 25, 1948, at the Meeting of the U. G. G. I. at Oslo, and on a manuscript Report, dated April 24, 1947. — The previous article I, in *Tellus* 1, N:o 1, is for shortness referred to as »I».

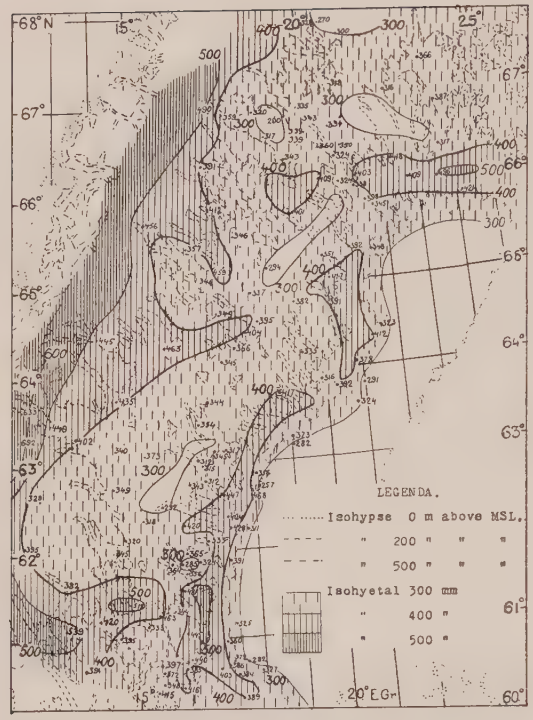


Fig. 1. Average precipitation (1909—1925) during August—April in N. Sweden (excepting the alpine region, the Scandes) showing a marked maximum zone along the Bothnian coast.

1. Rain-fall maps for the winter-half of the year, in temperate and polar regions, show coastal maxima of precipitation which lie even much nearer to the coast than one would expect from the classical theory of orographic precipitation.

(a) This fact is, for instance, revealed by the *rainfall distribution along the Bothnian coast of N. Sweden* in the season August—April, shown in fig. 1 (mean values for the years 1909—1925). E—W-profiles of the country at 61°, 63° and 65° N are seen in fig. 2, showing along the northernmost profile a gradual and small rise of the country from the very coast at 21°10′ E long. to about 16° E long. where the mountains begin. Nevertheless, even along this profile there is a marked zone of maximum precipitation (430 mm) less than 20 km from the coast, and a minimum (300 mm) in the somewhat higher interior of the country. For the two other profiles the corresponding figures are 420 and 290 mm, 510 and 325 respectively, showing an even more pronounced decrease of rainfall inland in spite of a general increase of the elevation of the country. — Most of this coastal precipitation is deposited with the upslope and/or convergent motion associated, in this region, with winds between S and NNE.

Not until near the watershed of the Scandinavian mountain range, the Scandes, and on its W. side, an amount of rainfall considerably greater than the Bothnian maxima is encountered. But then these great amounts are deposited by winds between SW and NW, i.e. during weather types which hardly contribute at all to the maxima along the Bothnian coast.

In the absence of the Scandes the winter precipitation deposited with the above-men-

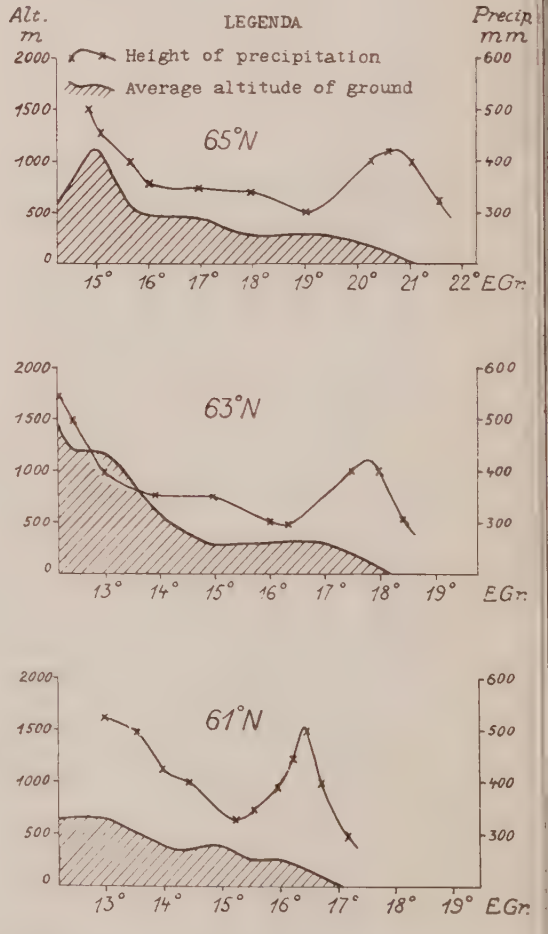


Fig. 2. Profiles of precipitation (from fig. 1) and ground from the Bothnian to the Scandes, along 65°, 63° and 61° N long., showing the marked coastal rainfall maxima.

tioned S to NNE winds over N. Sweden would obviously only show the marked Bothnian maxima of figs 1—2, and then a continued decrease of precipitation across the peninsula. An explanation of this, partly unexpected, distribution will be attempted in 3. But it should be stated already here that the discontinuity of friction, and of direct supply of water vapor into the air from the earth's surface, found along a coast-line evidently favours the formation of a zone of maximum precipitation near this line within air-currents with an on-shore component. If the ground is flat and comparatively dry the amount of precipitation in such currents will decrease more and more inland.

(b) To prove the validity of the above statements more in detail one has to resort to other methods than using only the official rainfall stations, cfr fig. 1, since this network is too sparse, as seen by that figure. Figs 3 A and B show *the snow-depth over the very flat Stockholm-peninsula*, at a map-scale 4 times that of fig. 1, during some well-defined weather-types in the cold winters 1941/42 and 1942/43. In both cases most of the snow fell during one week, without any thawing and with winds between E and N. The Baltic and Bothnian Seas were still mainly ice-free, and the moisture supply into the air, while passing these waters, was considerable. The snow-cover was insignificant at the beginning of each period, and the wind force was generally moderate when the snow fell. Considering that even this flat region in its NE. part is covered by forest to 65 % (excepting the outer skerries in the sea), and that the observers were instructed to measure the snow-depth, if possible, in forest glades, or in similar places, where it had not drifted, we may infer that the observations give at least a good relative measure of the amount of precipitation during these periods. Moreover, the agreement between neighbouring figures, and the close connection between the snow-depth and the orographic features (in this case mainly the distribution of sea and land) gives another, and even better, support to this statement.

The observers were mostly voluntary, and their cooperation was demanded by a broadcast appeal or by telephone calls to unknown people. The number of official rainfall stations in the 1-degree square 59—60° N lat. and 18—19° E Gr. long. is only 12, i.e. $^1/_8$ of the number of snow-depth data (98) within the same square in fig. 3 B. The official stations alone could have given little evidence concerning the problem discussed here, when used for a period of a few days only.

It is striking to notice that the special distribution of snow in fig. 3 A is repeated almost in all details in fig. 3 B, a year later. The weather type was in both cases characterized by a low ice-nucleus level and a low condensation level. The orographic cloud-system was as a whole stratiform, sometimes with a marked inversion above it already at about 1 km height. Owing to the intense heating of the cold air over the open sea, however, the air below the inversion was often unstable, and in reality the cloud-system consisted of very low and untypical, dwarflike, Cunb fused into a more or less continuous cloud deck. In such weather-types even a slight lifting, or an increase of turbulence, will presumably suffice to release convection, and formation of dwarf-Cunb, within the lowest unstable layer, at the coast-line. Farther inland this convection is probably damped or stopped by continental cooling, and in any case the moisture supply from below into the clouds is almost cut off. It is reasonable, then, to claim that any such orographic effects within the region investigated will be more pronounced than otherwise in winter time with an on-shore wind. — In fact, E of 17°30′ long. this region is remarkably flat, the isohypse of 60 m (200 ft) only comprising some very small areas at A, B and C on the maps, see fig. 3 A. Consequently, in this region any orographic effect, lifting the air, orderly or disorderly, considerably above the condensation level in winter, when the thermal convection due to the heating of land surfaces is missing, must be caused or released mainly by the frictional difference mentioned.

(c) A *direct frictional effect*, partly caused by the frictional convergence and lifting within an on-shore wind (Baldit 1929, Arakawa 1936), and partly by the sudden increase of vertical mixing at the coast-line, and affecting only a shallow layer of the atmosphere, needs a low condensation level to produce any cloud at all, and the lower the better. Even the upper limit of such a cloud will generally be low, and in order to get the water which has been condensed within it down to the earth again as

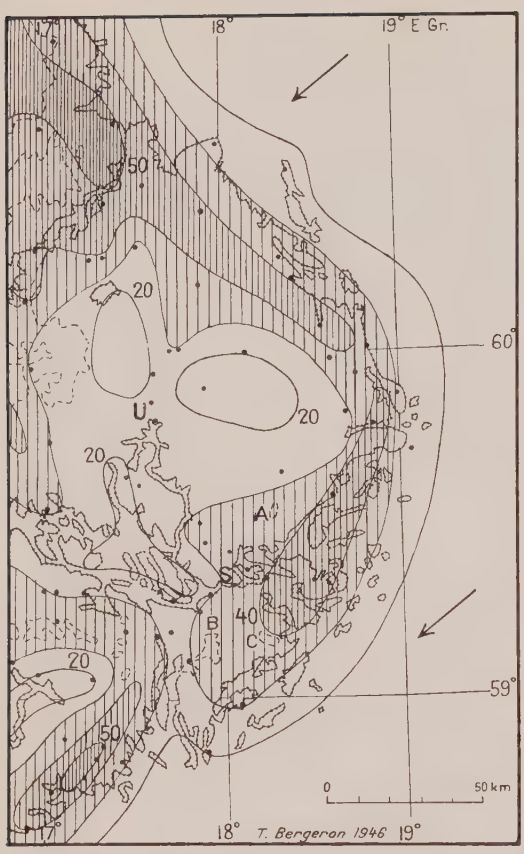


Fig. 3 a.

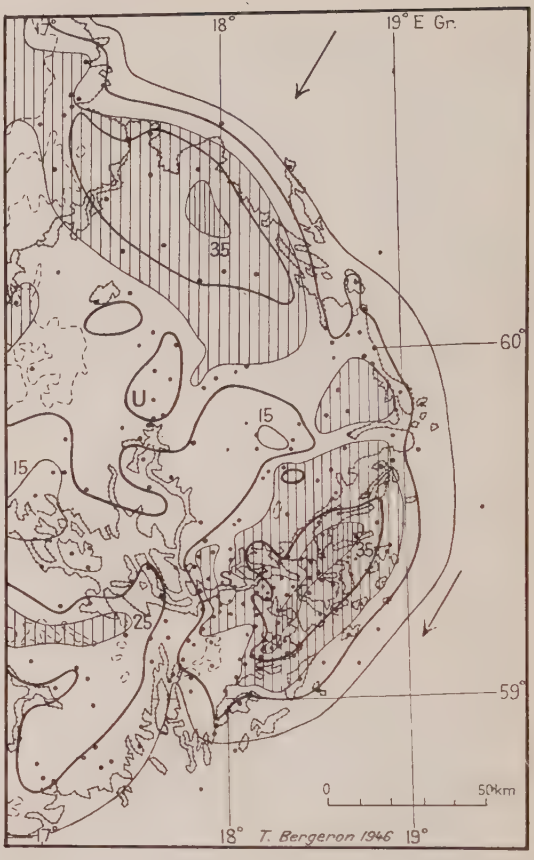
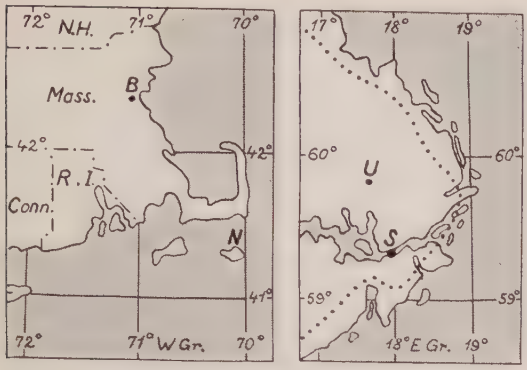


Fig. 3 b.

Fig. 3. Snow-depth distributions over the Stockholm-peninsula (region Uppland—Södertörn) resulting from snow-fall with NE wind (the arrows) in cold winter weather, showing combined frictional effects on precipitation along a flat coast.

rapidly and completely as possible the cloud must also become physically unstable, e. g. through infection by a suitable number of natural (or artificial) ice-nuclei (cfr 7., in article I). Thus, a low ice-nucleus level will mostly be another necessary condition for getting a marked coastal rainfall maximum along a flat coast, such as the one in figs 1—3, where friction is the main orographic factor. — Let us then state, however, that the winter conditions described by figs 3 A and B could be very little influenced by any artificial seeding, as the natural one was already operating, and probably nearly at an optimum rate. — In similar cases, with a much higher ice-nucleus level, not reaching down to the orographic coastal cloud, the precipitation might be considerably less intense, all other conditions remaining the same, cfr 3. Then, a suitable artificial seeding might increase the precipitation within the coastal zone of maximum condensation, but this would also mean depriving the current of some moisture which otherwise might give precipitation farther inland by a renewed orographic lifting, or by thermal convection.



B = Boston U = Uppsala
N = Nantucket S = Stockholm

Fig. 4. Dimensions of the region in fig. 3 compared with a region in New England, U. S. A.

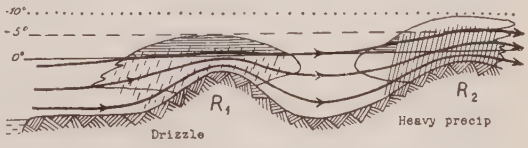
..... Central line of zone of maximum snow-depth in fig. 3 B.

The snow distribution in figs 3 A and B will be appraised anew in 3. after having learnt more about the orographic precipitation effects. — Before leaving figs 3 we shall compare the dimensions of their region with a well-known region in the USA: the SE. part of New England shown in fig. 4. Owing to the analogous orographic situation, and the frequent occurrence of blizzards with N-NE wind in New England, very pronounced coastal snow-maxima must sometimes form in this region, although, presumably, the network of rainfall stations will also here not be dense enough to show these maxima.

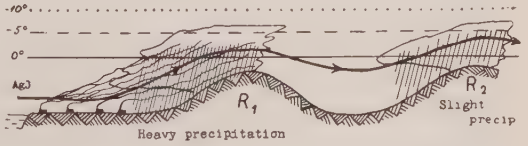
2. *The hydrometeoric reactions of on-shore and cross-mountain air-currents to fundamental types of orography*, partly discussed in 1., are represented schematically in fig. 5, showing land profile, air current, clouds, and precipitation in a vertical section from the sea (or very flat lowland) to the left and landward to the right. — The lefthand ridge R_1 may symbolize (1) a coastal mountain range (such as the Californian) or (2) simply the coastal effect of increased friction etc. discussed in 1., found along a flat coast (such as around Stockholm, or N of 64° N lat. at the Bothnian Sea). In both cases the on-shore wind component produces a cloud system. — The righthand ridge R_2 may symbolize (1) a (higher) inland range (such as the Sierra in California, or the Scandes in an easterly weather type), or (2), in summer, just the effect of increased thermal convection by the gradual heating and instabilization of the air farther inland. In both cases a second cloud system or cloud region may occur.

Fig. 5 A (upper picture) represents conditions in latitudes, seasons or weather types with a high 0°C-isotherm and ice-nucleus level (temperate latitudes in summer, warm spells during colder seasons, lower latitudes all the year), and natural seeding only. The coastal cloud is assumed to surpass the 0°C-isotherm but not to get any efficient precipitation release. It will then give only little precipitation, mainly as drizzle. — The inland cloud will reach the natural ice-nuclei or, within the Tropics, get efficient release even without their cooperation, cfr I, 3. Consequently, this cloud may give abundant rain provided that the condensation is intense.

Fig. 5 B (lower picture) represents (1) conditions in latitudes, seasons or weather types with a low 0°C-isotherm and ice-nucleus level such as discussed in 1., or (2) conditions of the opposite type A, but with a suitable artificial ice-nucleus seeding. It is easily seen that any artificial seeding (of the coastal cloud, this being the only unreleased cloud) has a similar effect as the transition from summer to winter or from a warm to a cold weather type. In both cases the only material result, if any, would be to transfer the zone of maximum precipitation from the (higher) inland districts to the coastal zone, where it is of little use as a power source, and where precipitation generally is sufficient anyhow. Moreover, in winter or in cold summer



A. With only natural ice-nucleus seeding (at R_2)



B. With efficient artificial seeding at R_1 (and R_2)

Fig. 5. Orographic cloud-systems produced in on-shore and cross-mountain flow (schematic vertical sections).
═══ Potential layer of precipitation release, cfr I, fig. 8.

weather an artificial seeding would have no appreciable effect at all.

There is still something to be said about the method to use in these cases for seeding the orographic cloud systems. Their precipitation intensity is small compared with the Cunb precipitation, and rainfall amounts of economical importance are obtained mainly thanks to the vast extent of this kind of cloud-systems and the long duration of their precipitation. The orographic cloud systems are stationary, and a prevailing wind is flowing through them. Consequently, any artificial seeding of them must also be of long duration since the nuclei are constantly brought down to the earth with precipitation elements, or blown out of the cloud-system to the leeward. Moreover, as stated in I, it is important to keep a suitable ratio v within the cloud to get the effect desired. These two conditions can hardly be fulfilled by dropping CO_2-pellets from air-planes; if practicable, this method would be uneconomic. The *quasi-permanent ice-nuclei* found by LANGMUIR and collaborators (AgJ and other chemical compounds) might possibly be used in this case since such nuclei can be produced at low cost in even a much greater number than those produced by the CO_2-pellets, and since they can be brought up into the clouds from below in the manner indicated in Fig. 5 B. A row of AgJ-factories, erected along a windward coast, are assumed to pollute the air with these nuclei — which can subsist also at temperatures above 0°C — and they are gradually spreading vertically through the whole depth of the air-current in question, and also horizontally, thanks to turbulent vertical and lateral mixing. The great problem, however, would be to keep a suitable v throughout all the orographic cloud-system, from its windward to its leeward edge, in view of the constantly increasing vertical spreading of the nuclei; cfr BERGERON 1943, p. 271—272.

Here we might, at the same time, see a glimpse of a solution to the main difficulty underlined above: the artificial seeding "stealing" rain from the dry inland instead of rendering more rain to these regions. If it could be possible artificially to keep the coastal cloud-systems more or less continually overseeded [cfr I, 7(b).] their release (by ice-nuclei) then and there might be impeded in all cases where this cloud system did reach the natural ice-nucleus level and otherwise would have rendered a rather abundant precipitation. At the same time more moisture would be left within the current for producing some possible precipitation inland where, as a rule, it is much more needed.

This method of overseeding might then also perhaps be used in some way or other to *prevent the too intense precipitation from, and electric phenomena within, quasistationary Cumulonimbi* by transforming most of their cloud droplets above the 0°C-isothern into minute ice-crystals. If artificial ice-nuclei were present in the same number as, or in a greater number than, the condensation nuclei ($v \leq 1$), and were of the same efficiency as the latter, only sublimation would occur above the 0°C-isotherm, since $E_i < E_w$. Then, any supercooled droplets brought up above this level by the intense updraft in the Cunb would be "consumed" by the crystals in a few seconds, cfr I, 7 (b). Such numerous ice-crystals cannot, any more than the ordinary cloud droplets of a Cu, grow very much by direct sublimation alone as long as they remain within their respective air parcels. Then, they will not acquire any appreciable fall velocity and the cloud will remain physically stable. The final result might be a cloud compound being a water-cloud, a Cumulus, below the 0°C-isotherm, and a kind of Cist above this level, giving no appreciable precipitation, especially no hail, no squall at the ground, and no lightning or icing in its interior — a tamed monster deprived of all its horrors.

The next step, to extend this treatment to the rapidly passing Cumulonimbi within a strong air-current, or to tropical hurricanes, does not seem very promising, in view of their propagation which mostly would make any stationary seeding plants useless. — *The dissipation of extensive storms by artificial seeding with ice-nuclei*, if feasible at all, could probably only be achieved by the aid of quasi-permanent ice-nuclei, as outlined above, and *not by dropping CO_2-pellets from air-planes*, as tried hitherto.

3. Synoptic investigations of certain weather situations, characterized by *excessive orographic precipitation*, carried through already many years ago, have led the writer to the conclusion, that *the pronounced coastal rainfall maxima*,

displayed by most winter-time rainfall maps, *are produced in a number of different ways.*

(a) Such maxima may occur even according to the classical scheme of orographic precipitation (by the orderly lifting of an on-shore and/or cross-mountain current) when the air is so humid that the condensation level, and thereby also the thickest cloud, comes to lie very near the sea, provided that there is an efficient precipitation release operating within this cloud.

(b) Other maxima are mainly conditioned by the direct frictional effects discussed in 1. above.

(c) A most striking case, the main features of which cannot be explained by the processes (a) or (b), is shown in Fig. 6. The general situation was governed by a steady outflow of continental Polar air, and partly also of Arctic air, from the Russian continent in connection with a quasi-stationary anticyclone over N. Scandinavia and N. Russia. A marked and quasistationary frontal zone (not shown in fig. 6) between the *cA* and *cP* to the NE and the *mP* (and partly *mT*) to the SW, extended from the Black Sea over N. Germany and S. Denmark to Scotland and Iceland, giving upslide precipitation on its northern side. The northern limit of this precipitation is represented by a zero-isohyetal running from Stockholm northwestward to Storlien and farther westwards over the Scandic, cfr fig. 6. The frontal precipitation measured during 3 days (24—27 March, 1927) increases from this zero-line towards the front zone and attains an average value over S. Sweden and Denmark of 20 to 30 mm. — A detailed study of the isohyetals reveals several striking deviations, though, from this general distribution, deviations which must be closely connected with the orography. The outstanding feature of the whole map is the narrow and very intense maximum zone along the Norwegian Skagerak coast, culminating at Grimstad, on the very coast, with 201 mm of melted precipitation in 72 hrs. (Most of this precipitation fell as snow or sleet, and in all it continued for about a week, with the effect that this small town, Grimstad, NW of **G** on the map, became literally buried in snow).

A zone of excessive precipitation may sometimes be caused by one giant Cumulonimbus, or by a so-called convective system.

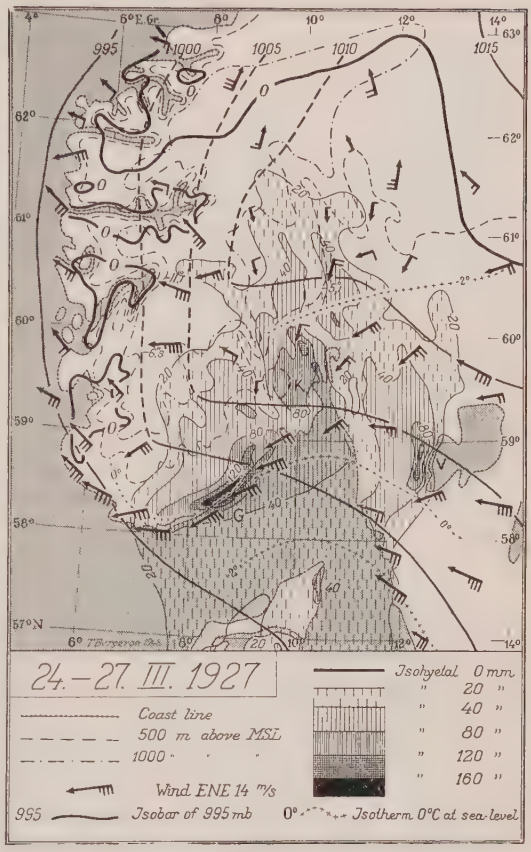


Fig. 6. Distribution of precipitation, wind, pressure and temperature in SW. Scandinavia during the period 24—27 March 1927 (72 hrs), showing marked coastal maxima of precipitation mainly due to orographically conditioned convergence.

In the case of fig. 6 such an interpretation, however, is first of all contradicted by the fact that the season and the general weather type hardly permitted the formation of any Cunb at all, not even of the dwarf-type discussed in 1., much less then of the giant-type sometimes forming inland in the summer-half of the year, and rendering 100—200 mm of rain in passing a region. The general current in lower layers was SE, at the mid-troposphere level presumably around W, and there was certainly rather warm air aloft above the front surface. The clue to the situation, and the final proof that this zone of excessive precipitation was of real orographic origin, is given by three facts:

(1) The same connection between orography

Table 1. Data from Norwegian mountain stations averaged for the period 24.—27. III. 1927 (07h—07h GMT).

Station	φ	λ	z(m)	T_z (°C) observed	T_0 (°C) interpolated	γ (°C/km)	DDF
Kutjern	60° 34′ N	10° 33′ E Gr.	496	— 5.1	— 2.5	— 5.2	NNE 4
Svandalsflona	59° 50′ »	6° 59′ » »	1 065	— 6.7	— 3.0	— 4.4	E 9
Slirå	60° 37′ »	7° 23′ » »	1 304	— 10.7	— 3.5	— 5.5	ESE 8

and precipitation is repeated, on a smaller scale, in two other regions of the map, viz. in the region of Kongsberg (**K** in fig. 6, W of Oslofjord), and just to the West of the lake Vänern (at 59°.0—58°.5 N and 12°.5 E Gr., **V** on the map) where a maximum precipitation of 108 mm was actually measured, whereas the precipitation from the same air-current over the Lake itself (at about 59° N, 13° E Gr.) was less than 15 mm.

(2) In these three regions, situated to the W of open water, the country is rising considerably and produces in all places a special wind pattern, treated below, at least within the lowest layers of the atmosphere.

(3) The orographic effects (1) and (2) were even found on every individual day during the period of fig. 6, and also during other analogous autumn and winter periods. Inger Bruun 1944 has several daily rainfall maps showing a quite similar distribution within Norway (cfr below p. 30).

The whole phenomenon is most pronounced in S. Norway, which has a more marked orography than the surroundings of Vänern. The former region, therefore, displays the effect in its typical form. Thanks to the vertical stability of the air over S. Norway during the period of fig. 6, shown by Table 1, the ESE current of the 0.5—1 km nearest to the ground could not surmount the Scandes in this region (where the mountain barrier is on an average 1 000—1 500 m high), but was divided into two branches. A minor branch found its outlet over the lower part of the Scandes near Storlien, flowing slowly upslope (Östersund and Röros 4 Beaufort), but quite rapidly downslope (e.g. up to 8 B West of Trondheim). The main current had to bend SW or even SSW (even Kutjern, at 500 m above s.-l., reports NNE 4 B) and to flow along the Skagerak coast as an increasing NE wind. This increase of the NE wind is due to its converging into the general ESE current coming over Sweden and Denmark, but the $\frac{\partial v}{\partial s}$ is evidently not great enough to compensate for the pure stream-line convergence, i.e. $\mathrm{div}_2\mathbf{v}$ remains < 0: *orographically conditioned convergence*. — The convergence is considerably accentuated by the fact that, in this case, the region of greatest roughness and friction (the land) lies on the high-pressure side of the current, an effect which has been recognized since 1919 both theoretically and empirically by the Bergen School: the *indirect frictional convergence* (cfr Bergeron 1934, fig. 51, p. 70). An intense lifting then must take place within this ENE current, although it is flowing almost parallel to the coast-line and to the general isohypses of the country; and this lifting is evidently reponsible for the excessive coastal snow fall — It should be noted here that these and other features of the wind-field seen in fig. 6 cannot be dismissed as occasional deviations, but are representative and persistent, since this field is constructed by forming the vector mean of 10 individual observations, at the main observation hours, through the whole period. [A number of other interesting features in the wind-field (not discussed here), in the pressure-field, and regarding their interaction, is reserved for future treatment.]

We now return for one moment to the *experimentum crucis*, the conditions along the W. shore of Vänern. Also here an open water surface borders to a rugged country, and the same orographically steered wind-pattern (see Vänersborg, at the southern part of the lake, with ENE wind), frictional difference, convergence, lifting, and cloud-system as at the Skagerak coast is produced, though on a smaller scale. — All the above reasoning does not, however, in itself explain definitely why the precipitation reached such extraordinary

amounts just in this March case. Favouring conditions were the strong wind and the negative values of $T - T_s$ over the water surfaces, both increasing the intensity of condensation within the coastal zone of maximum precipitation. — Some October cases, with the same general wind pattern, have also been studied. During the period 20—26 October 1933 the wind velocity was only about $^2/_3$ of that in the March case. On the other hand, T_0 was around + 10° C, and the condensation level was just as low, implying that the saturation vapour pressure E, and $\frac{dE}{dz}$, in an individual rising air parcel must have been even almost the double of the March values studied here. Nevertheless, in the October cases the greatest amount of precipitation in the coastal maximum zone was only 5—20 mm a day, and the precipitation was mainly drizzle. The latter circumstance helps us to solve the problem.

In fact, a cloud system formed by the orographically conditioned convergence discussed above, along a coast of only 100—500 m height generally does not reach the natural ice-nucleus level (at about — 10°) unless formed entirely within cA (which was not the case here), because as a rule the cloud-system will not extend much higher than the average height of the steering obstacle; above a certain height there will mostly even be a marked divergence, combined with a rapid decrease of the lifting motion and with vertical stability. The average mountain temperatures observed in the March case at the 500, 1 100 and 1 300 m levels (see Table 1 and fig. 6), indicate that even during this cold winter-situation the purely orographic cloud, at least in its southern part, around the sea-level isotherm of 0°C, did not reach the ice-nucleus level.

The processes prevailing during all such easterly weather-types in S. Norway could then be generalized as shown very schematically in fig. 7. The lines of equal values of negative $\mathrm{div}_2\mathbf{v}$ (convergence), of lifting, and of intensity of condensation to the East of the Scandic divide (and their counterpart with opposite signs on the W side) would then nearly coincide, any deviations being due to cross-current and down-stream differences in pressure, temperature and humidity.

In an imaginary case of this kind, without precipitation release, the water content of the

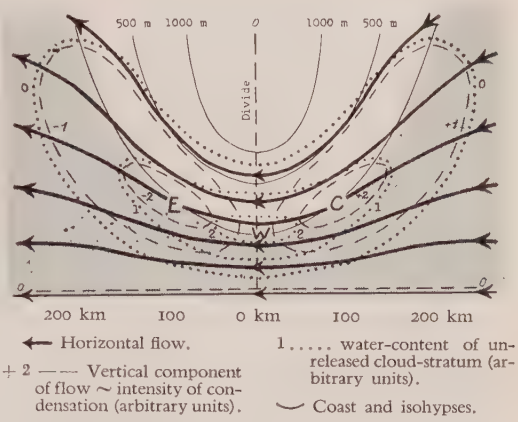


Fig. 7. Schematic map of orographically conditioned convergence, vertical motion, condensation and precipitation produced in a general E wind by a symmetric obstacle similar to S. Norway.

orographic cloud system would be strictly symmetrical with respect to the N—S line of divide — -- — in fig. 7, and could be represented by the dotted isopleths in this fig. — This state would be analogous to that of the stationary and symmetric cloud-cap of fig. 1 in article I, though the vertical axis of that fig., pointing upward, would correspond to the divide-line of fig. 7, pointing downward. Moreover, the horizontal dimensions are 50—100 times greater in fig. 7 than in fig. I, 1. Just as in the latter case, also within the orographic cloud-system of fig. 7 — if it were physically stable (I, 1.) — the individual cloud-droplets would rush through the cloud at high speed, forming and growing within the rising current at the Eastern (windward) side, shrinking and at last dissolving within the sinking current at the Western (leeward) side — and few of them would be deposited on the earth's surface.

In reality, however, there will always be some precipitation release within the cloud-system of fig. 7, thanks to the greater extent of this cloud-mass and the much longer lifetime of each droplet as compared with the local cloud-cap; cfr I, 8 (d). Under conditions such as in the October cases quoted above the release is inefficient and slow, though, so that only a fraction of the water condensed reaches the ground as precipitation; and this will not occur just below the place where the release started, but to the leeward. The

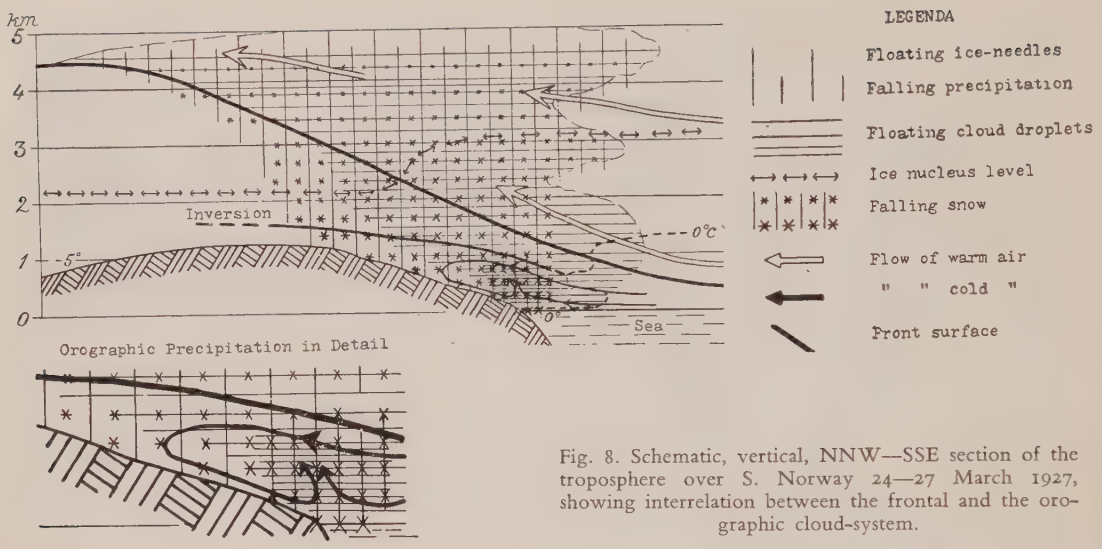


Fig. 8. Schematic, vertical, NNW—SSE section of the troposphere over S. Norway 24—27 March 1927, showing interrelation between the frontal and the orographic cloud-system.

maximum zone will then have a position in between the two positive maxima *C* and *W* in fig. 7, and the precipitation within it will be of modest amount and drizzly character, even if the intensity of condensation in the zone *C* is extraordinarily great.

In the case of 24—27 March 1927, shown in fig. 6, however, the synoptic maps (not reproduced here) clearly show that any low, orographic cloud systems forming within the SE current of Polar air at the SE coasts or SE slopes of S. Scandinavia were overshadowed by a vast and quasistationary shield of Altostratus-Nimbostratus, formed mainly within and above the afore-mentioned frontal zone. From this cloud-system snow fell continuously everywhere in the region between the zero-isohyetal of fig. 6 and the front-line south of Sweden (even in regions having orographic rain-shadow with an E-wind, except in the fjords of W. Norway, where the orographic shelter was too effective). North of the 0°C to + 1°C-isotherm at the ground (see fig. 6) the precipitation reached the ground as snow. Even south of this line snow must have reached down to the 200 or 300 m level, and thereby also to the tops of the orographic cloud systems of S. Sweden and Denmark, although these were not so strongly developed.

The three main zones of excessive orographic precipitation discussed here lie almost entirely within the first-mentioned region, with snow reaching the ground. In these zones the continuous snow-fall from a cloud having its base at 1 000—1 500 m above s.-l., as shown in fig. 8, represents a natural mechanism for ice-nucleus seeding and precipitation release even more effective than the ordinary one operating in supercooled water-clouds at and above the — 10° level. In this case the seeding elements (snow-flakes) were already large and fell at a rapid rate, thus forming a kind of "network" of branched crystals apt to carry down to the earth most of the water condensed in the low cloud within zone *C* of fig. 7, rapidly and almost directly below the spot where the condensation took place. Let us assume the fall velocity of the snow-flakes to be, say, 1 m/s, the centre of the main condensation to lie at 500 m above s.-l. (possibly bordered by an inversion at the upper limit of the ground turbulence and convection within the Polar air), and the average wind to be E 10 m/s through the lowest layer of 1 km depth (there will be a considerable decrease with height of this E wind above 500 m). A snow-flake would then be displaced only $\frac{500\text{ m}}{1\text{ m/s}} \cdot 10$ m/s, or 5 km, westward from its position at the level of main condensation, during its fall to the sea-level, a distance hardly discernible in figs 6 and 7. Under the assumptions made, it is thus permissible to claim that the area of maximum precipitation would

almost coincide with the area *C* in fig. 7. The shape and position of this area as compared with the area **G** in fig. 6 shows an agreement which is striking and to a certain extent convincing, even if we make allowance for the fact that the position and shape of the area *C* in fig. 7 is constructed only mentally, and qualitatively, from the stream-lines.

It is obvious that the zones **K** (W of Oslo) and **V** (W of Vänern) cannot compete with the zone **G** as to amount of precipitation and typical features, the orographic conditions being so much less pronounced. But they may well serve to confirm the validity of the general reasoning set forth here, because even snow-amounts of 100 mm (melted water) by *one* storm are very rare in Scandinavia. — As to the remaining zones of maximum precipitation during this period (only the one at Skagen is visible on fig. 6), none of them can serve as a further support of the existence of t h i s kind of orographic precipitation since there is nowhere else in the wind field any clear indication of the orographically conditioned coastal convergence. There is also nowhere else any i n d i r e c t frictional convergence effect acting since the wind nowhere blows with the sea to the left of the current (except just at the Southern coast of Scania).

On the other hand, most of these maximum zones are situated quite near the windward coast, even when the coast is very flat (Denmark, Swedish East coast between 55° and 59° N etc.), indicating that they are favoured or produced mainly by the d i r e c t friction effects discussed in 1(c).

4. Returning to the snow maps, figs 3 A—B, we may now state that only NE of Uppsala, where the wind was perpendicular to the coast, was the precipitation mainly produced by the direct friction effects. Within the maxima east and south of Stockholm, where the current was partly parallel to the general coast-line, the high-pressure side friction effect must have cooperated, and possibly also some orographically steered convergence at quasi-isohypsic motion. Only in that manner can we explain why these maximum zones are even more pronounced and their centers lie even nearer the coast than those NE of Uppsala.

It is important, therefore, to study the orographic precipitation in a completely flat region, where frictional differences alone can be made responsible for the distribution of rainfall. The Netherlands offer such an opportunity since the network of rainfall stations is tolerably dense (about 300), and since the NW. part of the country is extremely flat and even mostly depressed below sea-level, excepting the zone of dunes along its NW. sea-shore, see fig. 9. — A preliminary study of the rainfall distribution during individual days in July—October 1945 with extensive, and in part abundant, rain over the country has rendered a striking confirmation of the principles expounded above. This study is based on the 200 rainfall values printed *in extenso* in the Dutch year-book *Regenwaarnemingen*.

(a) Six cases, the 24^h-periods ending 1. VII., 18 ,19 and 24. IX., 24. and 26. X. 1945, with a S—SW wind, fresh to a gale on the coast, and rain from a frontal cloud system, gave in principle the same picture of the rainfall distribution, the map for 26. X. 1945, reproduced as fig. 9, being the most typical one. The outstanding feature of this map is the main maximum zone along all the sea-shore of Zuidholland and Noordholland, continuing toward NE and E over the Frisian islands, and practically coinciding with the belt of dunes, which forms the border of the country toward the North Sea and represents a very narrow (0—5 km) barrier of an average height of 20 m (5—60 m). Already C. Braak 1933 ascribes the marked maximum of average October precipitation found over Noordholland, well to the E of the dunes (l. c., Plate V), "to the friction above the land, and particularly by the dunes" (l.c., p. 100). In view of the much closer relation between the coast and the maximum zone during the special 24^h-periods investigated here, it is natural to go some steps further and interpret these maxima as analogous to those of figs 3, the one over the mainland forming with an on-shore wind component and corresponding to the northern maximum zone of figs 3, the one over the Frisian islands corresponding to the southern maximum of figs 3.

This Dutch analogy is especially valuable since it occurs with the opposite wind direction (warm-air current from SW) as compared with the Swedish winter cases (cold-air current from NE), but still (1) occurring with the greater friction on the high-pressure side

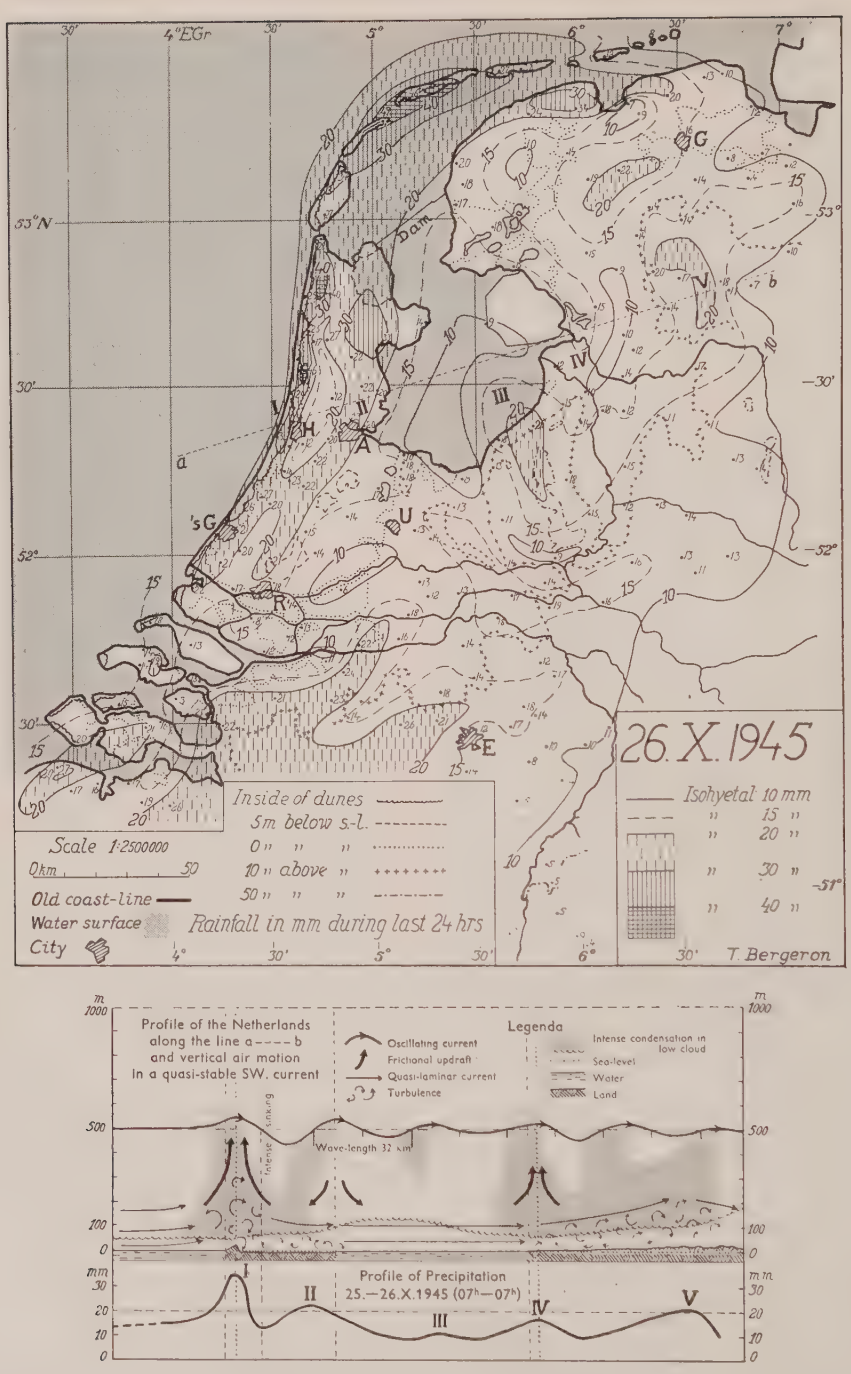


Fig. 9. Distribution of rainfall over the Netherlands, and WSW—ENE profile around 52.5° N lat., during the 24h-period ending 26. X. 1945, 07 GMT.

Map figures show the total rainfall in mm
a---b shows the position of the profile
A = Amsterdam 's G = 's Gravenhage
E = Eindhoven H = Haarlem
G = Groningen R = Rotterdam
U = Utrecht

of the current, hinting that this latter condition is a necessary, and even deciding, one for producing a typical and marked coastal maximum. Again, both in the Swedish and Dutch cases (2) the condensation level was low, and (3) the statical stability was pronounced, at least above the turbulence layer, hinting two other necessary conditions. Lastly, in *all* cases (4) the precipitation release was efficient, in the cases of figs 3 thanks to low temperatures at all levels, in the case of fig. 6, and in the Dutch cases, mostly thanks to the precipitation from overlying frontal clouds (cfr fig. 8).

(b) In fact, two other cases of abundant precipitation in the Netherlands from the same year, when one or more of the above conditions were failing, show quite different rainfall distributions. — During the 24^h-period ending 10. VIII. 1945 abundant rain fell over the country in connection with an upslide front to the E, belonging to a small retrograde cyclone over W. Germany, and with low level winds around NE in the province Groningen, but elsewhere with winds between N and WNW on the coast. In this case the main rainfall maximum (about 70 mm) lay near Groningen (coinciding with a seasonal maximum, cfr BRAAK, l. c., plate IV). Another maximum (55 mm) lay on the northern slopes of the Veluwe (max. height 107 m a. s.), in good agreement with the orographic effect to be expected in a WNWly current. — On the coast there was only an ill-defined secondary maximum (35 mm), in the provinces Zuidholland and Zeeland, produced in a Nly—NWly wind by the direct friction effects, cfr 1 (c), but counteracted by the indirect one, factor (1), cfr 3 (c). Moreover the former factors were not pronounced since the Nly current formed a comparatively unstable cold-mass. Only in Zuidholland—Zeeland the direct frictional effects were operating, presumably assisted by efficient release from the overlying frontal cloud, factor (4). — On the 26. X. 1945, some time after midnight (i. e. a few hours before the 24^h-period immediately following the period of the map, fig. 9), a postfrontal unstable cold-mass (mP) invaded the country from WSW (the coastal winds changed from SW'S to WSW), and at 3—8 km height above Uccle the temperatures fell between 4° and 10°C, whereas they remained almost constant in the lowest layers of the region. This implies the failing of the factors (2) and (3) above, whereas the precipitation release (4) may still have been efficient within the convective clouds. The result was a precipitation distribution differing entirely from that of the previous day, without any coastal maxima at all in Holland and Friesland, but with one maximum of 50 mm over the province of Groningen, factors (1) and (4), and another one of 50 mm over the river region.

The map of fig. 9 shows also some other noteworthy features. The maximum zone I, discussed above, is flanked to the E by a marked minimum zone at a distance of 6—10 km inland over Noordholland, i.e. over a very flat depression (— 1 to — 5 m), having in some parts only about 25 % of the rainfall of the maximum zone. In Friesland there is not always a clear indication of a minimum over the sea surface between the isles and the mainland — as could be expected. The weak secondary maximum on the N coast of that mainland (Friesland) could be expected from the indirect frictional effect, just as the secondary maximum (15—18 mm) on the E coast of IJsselmeer IV must be due mainly to the direct frictional effects. — A most surprising feature, however, is *the marked secondary maximum II over the E part of the absolutely flat land-surface of Noordholland* (faintly indicated also in Zuidholland). In the profile at the bottom of fig. 9 — drawn along the line a - - - - b in the map, and parallel to the gradient wind — this maximum is explained as caused, or at least supported, by the first free oscillation of the low troposphere after the obstacle formed by the dunes and through the considerable increase of turbulence in the lowest layers set up by them. It is reasonable to compare this effect with the effect of a smooth obstacle in a laminary current of at least 200 m height. According to recent investigations and experiences in connection with soaring flight, such an obstacle may cause marked oscillations several km up in the atmosphere; cfr KÜTTNER, 1939, p. 113, fig. 43 and QUENEY 1948, p. 19, fig. 1. The next maximum (III) is partly conditioned by the Veluwe. It is possible, though, that this updraft, and the updrafts causing the faint Overijssel rainfall maxima (IV and V), as a

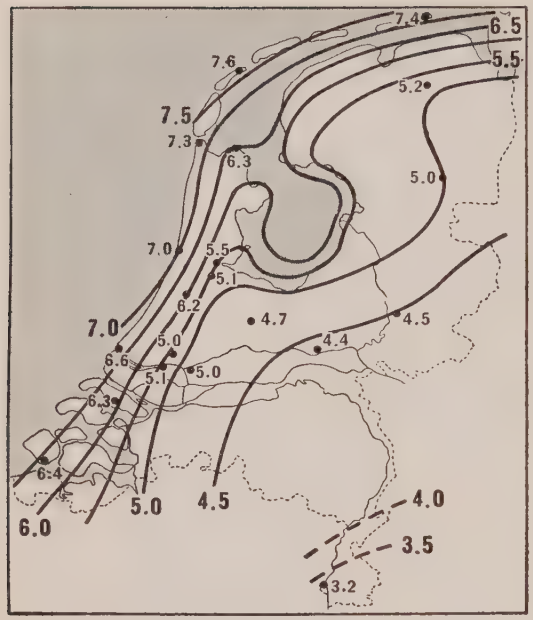


Fig. 10. Average wind velocity (m/s) in the Netherlands in winter (Dec.—Febr.). — From C. Braak 1942.

component contain the second, third and fourth free oscillations after the main obstacle at the dunes. Moreover, it seems reasonable to assume that the asymmetric position of maximum II (E. Noordholland) is caused by frictional downdraft at the W. shore of IJsselmeer, see profile of fig. 9. In fact, the climatological wind maps prepared by C. Braak 1942, show a very marked influence on the wind even of this flat country (see fig. 10), the average velocity in Dec.—Febr. decreasing from 7 m/s on the W coast to 5 m/s in the interior of Noordholland, and again attaining almost 7 m/s over the IJsselmeer.

Naturally, the question arises to what extent these different maxima are able to leave their marks on the average (normal) rainfall distribution of this region and season. In fact, the Dutch normal rainfall maps edited by C. Braak 1933 show one astonishingly pronounced secondary maximum over Noordholland in all the months August to December, whereas it seems to fail entirely in the months March—July: conditions (1) and (3) may fail completely then! As aforesaid, Braak attributes the just mentioned maximum to frictional differences. On the other hand, there is no trace, on these normal maps, of double maxima in Noordholland — and it would be unreasonable to expect them. Their intensity and position with respect to the dunes, and more particularly their mutual distance, will vary considerably even with small variations of the direction, velocity, humidity and statical stability of the general S—WSW flow producing them. Consequently, say, during a rainy October month with a southwesterly weather type, two such maxima, even if they coexist as separate maxima during each single day, will overlap so frequently that the average picture will be blurred and only show one single maximum. Moreover, also during such a month there will always be many days with situations like those described in (b) above which will blurr the picture even more.

In this connection the author should like to emphasize the main difficulty felt by all who have tried to study different precipitation distributions and factors: *the period of* 24^h *between two consecutive measurements is too long.* In most cases it is impossible to find — over a region of some extent — a 24^h-period containing only a "one-factor rain". At the beginning or end of the period other factors generally will creep in and disturb the results of the study. The chance of catching a "one-factor rain" over a region would become more than doubled, if the period was cut down to 12^h, as is the case with the stations reporting by telecommunication. Their network, however, is much too sparse for the purpose in view, cfr 1 (b). Consequently, *it would be desirable to introduce twice-daily measurement at all rainfall stations* in some countries, at least during certain months or years.

5. The *effect of artificial seeding* on an orographic cloud system of the types discussed here may now be expressed in short terms.

(a) If a frontal cloud-system is producing a natural seeding of an orographic one, as in the case of March 1927, then an artificial seeding has no chance of increasing its precipitation. The only possible artificial change within the cloud-systems might be achieved by overseeding the frontal cloud system, thereby impeding its precipitation, which in turn would check precipitation from the (coastal) orographic one.

(b) The important points with the orographic cloud-systems are: (1) that a number of them may just surpass the 0°C-isotherm, but not reach the natural ice-nucleus level, (2) that the intensity of condensation may be considerable within them, and (3) that the above conditions may be stationary for hours or days.

(c) In the cases (b), with a slow and inefficient natural release, as in the October cases of 1933, the artificial seeding might evidently bring about a big and steady increase of the coastal precipitation. These cases will mainly occur in spring and autumn in our latitudes. — Cfr, however, what was said in 2. about the little practical use of increasing coastal precipitation.

6. We cannot leave the question of orographic zones of maximum precipitation without mentioning the problem of "*the increase of precipitation with height*". After having hinted, and partly studied in some detail, in articles I and II, precipitation mechanisms of widely different structure and functioning, it seems futile to lay down rules or to make quite general statistics regarding this assumed "increase" taken as a geographical phenomenon.

Obviously, almost every kind of precipitation mechanism will show an increase of precipitation from the ground up to the condensation level or cloud base of that individual mechanism thanks to the gradual evaporation of the precipitation elements below that level (cfr I, 3. and 12., and FINDEISEN 1939, p. 457). — With mechanisms having a high cloud base, well above all the meteorological stations (or mountains) of a country, such as the pure frontal Ast-Nbst-system or the Cunb, abstracting from any low (orographic) cloud underneath them, precipitation will increase with height, except in the rather frequent cases when these mechanismes themselves deform, weaken, or get impeded in their development or progress also by orographic barriers lying below the cloud base in question. — There is, for instance, much evidence in favour of a decrease also of convective precipitation toward the interior of high plateaus even if they lie near water, have rich vegetation and are snow-free. — With precipitation mechanisms having a low cloud base, the rule of increase with height will only hold good over comparatively low land.

Even greater complications meet in the case of precipitation mechanisms such as those studied in 1.—2., since they do not form or progress into the highland, some of them not even entering the low inland. — Any such statistics, therefore, should only be carried through for separate seasons, and for well defined homologous geographical regions and weather types.

7. It is appropriate to conclude this treatment of orographic precipitation by discussing some earlier monographies and other papers on such rainfall in NW. Europe in the light of the above findings. Most of them, however, treat the problem merely descriptively, without trying to solve the problem physically or to separate the different orographic factors operating. — The first noteworthy work of this kind for Sweden was achieved by C. J. ÖSTMAN 1927, who studied the distribution of cyclonic or frontal precipitation in S. Sweden with the four main wind directions SW, NW, NE and SE. For each wind, his maps show rainfall maxima on the windward slope of the central highland of S. Sweden (up to 370 m a. s.), and of the foothills of the Scandes, beginning at about 60° N lat. — as could be expected. Now we may also state that, with a NE and SE wind, these maxima are mostly located in accordance with the effects discussed above in 3. (cfr l. c., fig. 4 and 5). — For Norway the pioneer work in this field was performed by F. SPINNANGR through a series of papers (in *Naturen* 1925 ff.) on the rainfall distribution in S. Norway at different weather types. The results as to frontal precipitation are summed in his paper 1943. With NE wind below an upslide surface his maps show, over SE. Norway, in principle the same rainfall distribution as fig. 6 in the present paper, with marked coastal maxima and the amount decreasing landinward, the isohyetals being on the whole parallel to the isohypses, but *precipitation decreasing with height* (cfr l. c., p. 457, fig. 7). With S—SW wind, below an upslide surface, the analogous distribution is clearly seen on the Western coast of Norway (e.g. in fig. 8, l. c., p. 461), although the maxima in some cases lie up to 50 km inland from the coast. Also here the precipitation goes down to relatively small

values even far W of the great divide in the Scandes — contrary to any theories of a general increase of precipitation with height. SPINNANGR's results were corroborated as to all essential features, for SE. Norway, by the papers of INGER (HALVORSEN) BRUUN in 1943 and 1944. In the former paper, for the first time, we meet a clear statement of the importance of "the wind field in advance of the front" (i.e., in the terminology used here: the wind direction within the cold wedge below the upslide surface) for determining the rainfall distribution, although without any further explanation. 9 out of her 10 cases published in 1944 show in principle the same wind and rainfall distribution as fig. 6 above; even the bipartition of the marked maximum along the Skagerak coast is clearly seen. INGER BRUUN supposes these cases to represent frontal precipitation at all possible orientations of the front and with widely different weather situations resp. wind directions. In reality, however, the similarity of the rainfall patterns in all these cases hints that the bulk of precipitation fell prefrontally and within an NEly—SEly cold-air current (although this cannot be proved, because of the too great length, 24 hrs, of the period of measurement), essentially of the same structure as in the case of 24—27 March 1927 investigated above. In fact, this is the only weather type giving widespread excessive precipitation on that coast in autumn and winter. — It follows, then, hat when excessive rainfall occurs in connection with warm-fronts, this rainfall will often not be of true frontal or upslide character, but only conditioned by the front: a special low rain mechanism is formed by the cooperation of prefrontal wind-field, orography and the overshadowing frontal cloud (mainly rendering the efficient natural seeding). — In 1947 Sv. JOHANSEN delivered the investigation for N. Norway corresponding to those of SPINNANGR and INGER BRUUN. His maps confirm on the whole the action of the same orographic factors as in S. Norway. Because of the scarce network of stations JOHANSENS maps cannot be quite so convincing as those of SPINNANGR and INGER BRUUN. JOHANSEN's fig. 8 a—d (l. c., p. 368—369), though, shows a very marked prefrontal coastal rain maximum (40 mm) ahead of a warm-front, advancing from the NE, between 21° and 29° E Gr. long. where the prefrontal N—NW wind had a chance of flowing with land on its high-pressure side; the opposite was the case to the W of 20° E, and here the coastal maxima no longer were pronounced. With a SW wind parallel to the coast, on the other hand, the coastal maxima in this latter region become very pronounced. — C. L. GODSKE 1947 utilizes a very dense network of rainfall stations, erected for special purposes in Hardanger-fjord and around Bergen, for studying, on the basis of the classical theory, merely the huge differences in rainfall at short distances due to a very marked alpine orography. VEIKKO ROSSI 1948 studies the general effect of the Scandes on the rainfall distribution over all Fennoscandia in connection with fronts advancing from the W. The network of stations used is too scarce and the map-scale too small to permit any conclusions regarding special coastal precipitation effects.

8. Already in 1921 J. BJERKNES and H. SOLBERG made an important contribution to the physical understanding of orographic precipitation: the difference of friction between sea and land might cause convergence, lifting of the air, and precipitation. At the same time, however, they claim (l. c., p. 9) that only mountains of 1 000 m height or more can produce pure orographic rain, a statement which seemingly contradicts the former one.

Further, they discuss the effects of orography on fronts. This line of thought was pursued by T. BERGERON 1934 (l. c., §§ 34—35) and 1937, by investigating the effect of orography also on the field of wind and precipitation (cfr p. 6 above) in different air masses. — We have already quoted C. BRAAK concerning the outstanding importance of friction for explaining the distribution of the autumn and winter rains in the province of Holland. — In 1936 R. M. POULTER studied configuration, air mass and rainfall, and tried to supplement the classical theories for estimating the amounts of orographic rainfall by introducing the concept of the *front* into the calculations. According to his theory a warm-front would, for instance, give about the double amount of rainfall when approaching a hill-slope of the same inclination. We have tried above to show that orography only indirectly increases frontal rain, or *vice versa*, and merely by means of quite special mechanisms the yield of

which cannot be calculated in a simple way. — In 1945 L. C. W. BONACINA discussed the global aspects of orographic rainfall and maintained that such rainfall will as a rule have an optimum yield with mountain ranges of about 3 000 m height, but that even hills of 100 m height may cause such rain. Our rainfall maps of Sweden and the Netherlands surely point to much lower limits in both these cases, at least in NW. Europe in autumn and winter. BONACINA also emphasizes the important fact — first mentioned by the Bergen School — that orography may impede the progress of rain-producing mechanisms and thereby increase the rainfall in the mountains, and especially on their windward slopes.

The most important contribution in late years to the field discussed here is due to C. K. M. DOUGLAS and J. GLASSPOOLE 1947. Their paper is really elucidating as to the possible causes of orographic precipitation in certain weather types. Space does not allow us to discuss thoroughly the coincidence and differences between their results and those of the present paper. They have mainly studied orographic precipitation with strong winds between S and W in the rugged country of NW. Great Britain, conditions which differ widely from those represented by figs 6 and 9 above. We shall only quote one of their main findings: the greatest amounts of orographic rainfall often occur with a deep moist layer of tropical air blowing for a long time within a wide warm sector. The discussion in 4. above indicates that this situation may produce, and for a long time maintain, low-level orographic cloud-systems over the mountains in the NW, with intense condensation (high wind and temperature, statical stability, low condensation level, great depth of cloud). In spite of the high temperature such a cloud-system will receive efficient precipitation release from its upper part or from general cyclonic clouds thanks to the great depth of the moist layer. — One cannot think of any other southerly to westerly weather type in Great Britain that so well could fulfil these conditions, corresponding to the factors (1)—(4) in 4. It remains to see, though, how and where excessive orographic rain forms in the British Isles with a northeasterly weather type.

REFERENCES

ARAKAWA, H. 1936: The Effect of Coasts on Winds. *Geoph. Mag. Tokyo*, **11**, No. 3.

BALDIT, ALBERT 1929: Météorologie du relief terrestre. Vents et nuages. Paris 1929.

BERGERON, TOR 1934: Trechmerno-svyasnoy sinoptichesky analis, II. Z.U.E.G.M.S. Moscow 1934.

BERGERON, TOR 1937 (English summary of previous item): On the Physics of Fronts. *Bull. Am. Met. Soc.*, **18**, No. 9.

BERGERON, TOR 1943: On some Meteorological Conditions for the Dissemination of Spores, Pollen etc., and a supposed Wind Transport of Aloina Spores from the Region of Lower Yenisey to South-Western Finland in Juli 1936. *Svensk Botanisk Tidskrift*, **38**, No. 3.

BERGERON, TOR 1947: Report on a discussion with Dr. IRVING LANGMUIR on the problem of artificial release of snow within the atmosphere. Dated Cambridge, Mass., April 24, 1947 (Manuscript).

BERGERON, TOR 1949: The Problem of Artificial Control of Rainfall on the Globe, I. General Effects of Ice-Nuclei in Clouds. *Tellus*, **1**, No. 1.

BJERKNES, J. and H. SOLBERG 1921: Meteorological Conditions for the Formation of Rain. *Geof. Publ.*, **2**, No. 3.

BONACINA, L. C. W. 1945: Orographical Rainfall and its place in the Hydrology of the Globe. *Q. J. Met. Soc.*, **71**, No. 1.

BRAAK, C. 1933: Het Klimat van Nederland. A. Neerslag Erste gedeelte. Kon. Ned. Met. Inst. No. 102, *Mededeelingen en Verhandelingen*, **34 a**.

BRAAK, C. 1942: Het Klimat van Nederland. D. Wind. Kon. Ned. Met. Inst. No. 102, *Mededeelingen en Verhandelingen*, **46**.

BRUUN, INGER (HALVORSEN) 1944: Further Synoptic Studies on the Distribution of Precipitation in South-Eastern Norway. *Met. Ann.*, **2**, No. 2.

Douglas, C. K. M. and J. Glasspoole 1947: Orographic Rain in the British Isles. *Q. J. Met. Soc.*, **73**, No. 1.

Findeisen, Walter 1939: Das Verdampfen der Wolken- und Regentropfen. *Met. Zeitschr.*, **56**, pp. 453—460.

Godske, C. L. 1947: Nedbör, lys og varme i Sörfjordsbygdene. Festskrift for Ullensvangs Hagebrukslag. Bergen 1947.

Halvorsen, Inger 1943: On the Distribution of Precipitation over South-Eastern Norway when Fronts pass from the West and from the South-West. *Met. Ann.*, **1**. No. 11.

Johansen, Sverre 1947: On the Distribution of Precipitation in Northern Norway in Various Weather Situations. *Met. Ann.*, **2**. No. 11.

Küttner, J. 1939: Moazagotl und Föhnwelle. *Beitr. z. Phys. d. fr. Atm.*, **25**, pp. 79—114.

Poulter, R. M. 1936: Configuration, air mass and rainfall. *Q. J. Met. Soc.*, **62**, No. 1.

Queney, Paul 1948: The Problem of Air Flow over Mountains: A Summary of Theoretical Studies. *Bull. Am. Met. Soc.*, **29**, No 1.

Rossi, Veikko 1948: On the Effect of the Scandinavian Mountains on the Precipitation Fronts approaching from the Sea. *Fennia*, **70**.

Spinnangr, Finn 1943: Synoptic Studies in Southern Norway. II. Front Precipitation. *Met. Ann.*, **1**, No. 17.

Östman, C. J. 1927: Studier över nederbördens fördelning vid olika vindar i Svea- och Götaland. *Medd. fr. Stat. Met.-Hydr. Anst.*, **4**, No 2.

Long Waves and Cyclone Waves

By E. T. EADY, Imperial College of Science, London

(Manuscript received 28 Febr. 1949)

Abstract

By obtaining complete solutions, satisfying all the relevant simultaneous differential equations and boundary conditions, representing small disturbances of simple states of steady baroclinic large-scale atmospheric motion it is shown that these simple states of motion are almost invariably unstable. An arbitrary disturbance (corresponding to some inhomogeneity of an actual system) may be regarded as analysed into "components" of a certain simple type, some of which grow exponentially with time. In all the cases examined there exists one particular component which grows faster than any other. It is shown how, by a process analogous to "natural selection", this component becomes dominant in that almost any disturbance tends eventually to a definite size, structure and growth-rate (and to a characteristic life-history after the disturbance has ceased to be "small"), which depends only on the broad characteristics of the initial (unperturbed) system. The characteristic disturbances (forms of breakdown) of certain types of initial system (approximating to those observed in practice) are identified as the ideal forms of the observed cyclone waves and long waves of middle and high latitudes. The implications regarding the ultimate limitations of weather forecasting are discussed.

The present paper aims at developing from first principles a quantitative theory of the initial stages of development of wave-cyclones and long waves. For reasons of space and readability both the argument and the mathematics have been rather heavily compressed. A fuller and extended treatment of several of the points raised will be given in subsequent papers.

I. The Equations of Motion

Owing to the complexity (and non-linearity) of the simultaneous partial differential equations governing atmospheric motion it is desirable to simplify these by the omission of all those terms which do not make a major contribution *to the particular type and scale of motion envisaged*. This procedure is made possible by the fact that we know, from observation, roughly what the answers must look like. Its utility is justified by the fact that, having once obtained a crude model of the motion, we may then by successive approximation take into account any or all of the terms originally omitted. In the present instance we are concerned with relatively rapid development, by comparison with which radiative processes (or rather their differential effects) are slow. For a first approximation therefore we consider the motion as adiabatic. Also we are concerned with the motion of deep layers and for a first approximation we neglect the effects of internal friction ("turbulence") and skin friction. A rough calculation shows that the energy dissipated in the surface friction layer is usually much less than the energy supply to the growing disturbance and this is probably, in most cases, the major source of energy loss. The present paper is concerned only with systems which are initially (and also after a small perturbation) convectively stable, i.e., with those systems which would appear to be least favourable with regard to instability. Hence we use a system of equations appropriate to laminar frictionless adia-

batic motion of a rotating baroclinic fluid. Restriction to convective stability makes possible a further slight simplification in that in almost all[1] such cases (except very close to the border-line) we may neglect vertical accelerations and use the hydrostatic equation — the disturbances are "quasi-static". Briefly, the explanation is that in such cases the energy associated with horizontal perturbations greatly exceeds that associated with vertical motion.

The atmosphere is a compressible fluid and in estimating the significance of this fact it is convenient to consider separately the static effect, manifested by the decrease of density with height, and the dynamic effect, manifested by elastic forces in the equations of motion. The static effect has two consequences. In the first place, the static stability, a measure of the force tending to restore a displaced particle to its equilibrium position, is measured not by the vertical density gradient but by the gradient of potential density (or by the difference between actual and adiabatic lapse-rate). As compared with incompressible flow this involves only the modification of a parameter. In the second place, a given mass of air occupies, at higher levels, a greater height range. The result is that atmospheric flow can never be quite equivalent to incompressible flow but (as may be inferred from the detailed treatment) the difference is to be regarded as a distortion rather than any difference in kind. For systems which are not too deep there is an equivalent incompressible system whose behaviour closely parallels that of the atmospheric one. The nature of the "correction" for very deep systems is discussed below. The significance of the elastic forces depends on the type of solution in which we are interested. In the theory of atmospheric tides and the diurnal variation of pressure, where the wave-velocities are comparable with the speed of sound, these forces play an essential part. But in all waves associated with "weather" the wave-velocities are, by observation, small compared with the speed of sound and this is true, as we shall see, of both the real and imaginary parts of the wave-velocities of the theoretical solutions. We are therefore justified in treating the motion, from the dynamic point of view, as incompressible. The net result is that we can construct an equivalent "incompressible flow" problem and then use as our continuity equation:[1]

$$\mathrm{div}_H \mathbf{v} + \frac{\partial}{\partial z} V_z = 0 \qquad (1)$$

For systems which are not too large we may use the ordinary cartesian co-ordinates fixed in the earth in which we imagine a small spherical cap to be "flattened" on to the tangent plane so that gravity acts along parallel lines. This involves a certain amount of geometrical distortion (and consequently a distortion of our solutions). A more serious error results from the assumption of a constant Coriolis parameter and we may obtain a first approximation to this error by using the same co-ordinate system but regarding the Coriolis parameter as a function of y (the N—S co-ordinate). A more precise treatment of long waves requires the use of a polar (or equivalent) co-ordinate system and crude solutions, using numerical methods, have been obtained in this case. The broad resemblance of these solutions to those obtained by analytical methods in the cartesian system justifies the use of the latter as a first rough approximation.

If we define, for unsaturated air:

$$\Phi = \frac{1}{\gamma} \log p - \log \varrho \qquad (2)$$

so that Φ is proportional to the entropy, we have for adiabatic motion:

$$\frac{d}{dt} \Phi = 0 \qquad (3)$$

We shall *define* the static stability as $\frac{\partial}{\partial z} \Phi$. To study the motion of saturated air in contact with cloud we have only to alter the effective value of the static stability (our norm is now the wet adiabatic). Usually the reduction is

[1] The exception is the case of strong anticyclonic horizontal shear approaching in magnitude the Coriolis parameter—see below.

[1] When not otherwise stated, the symbols employed are those normally used in theoretical meteorological literature.

very significant, the forces opposing vertical motion being much less inside than outside a cloud, a fact having, as we shall see, important consequences. Moreover we obtain a direct translation from atmospheric motion to incompressible flow by the substitution:

$$\Phi \to -\log \varrho \tag{4}$$

Alternatively, we may regard γ as having, for saturated air, a value slowly varying from considerably less than 1.4 at high temperatures but asymptotically approaching the dry adiabatic value at low temperatures. For an incompressible fluid γ is effectively infinite.

With the approximations mentioned above we have the dynamic equations (where K is the Coriolis parameter):

$$\begin{aligned} -\frac{1}{\varrho}\frac{\partial p}{\partial x} &= \frac{d}{dt}V_x - KV_y \\ -\frac{1}{\varrho}\frac{\partial p}{\partial y} &= \frac{d}{dt}V_y + KV_x \\ -\frac{1}{\varrho}\frac{\partial p}{\partial z} &= g \end{aligned} \tag{5}$$

These equations, together with equations (I. 1)—(I. 3), form a complete set. Further simplification is however possible. (For the present we take K to be constant.) By partial differentiation with respect to z of the first two of equations (I. 5), using equation (I. 2) and the last of equations (I. 5):

$$\begin{aligned} \left(\frac{\partial}{\partial z} - \frac{\partial \Phi}{\partial z}\right)\left(\frac{d}{dt}V_x - KV_y\right) &= -g\frac{\partial \Phi}{\partial x}, \\ \left(\frac{\partial}{\partial z} - \frac{\partial \Phi}{\partial z}\right)\left(\frac{d}{dt}V_y + KV_x\right) &= -g\frac{\partial \Phi}{\partial y}, \end{aligned} \tag{6}$$

where the first brackets are regarded as operators. We shall see that the solutions with which we are concerned "oscillate" in the direction of z (it is immaterial that the "oscillation" is not sinusoidal) and if this "oscillation" is sufficiently rapid (if the disturbances are not too deep) the effect of the operator $\frac{\partial}{\partial z}$ swamps that of the multiplier $\frac{\partial \Phi}{\partial z}$. By numerical substitution in the final answers we can verify that this is the case for the solutions in which we are interested. Then for a first approximation we may use in place of (I. 6):

$$\begin{aligned} \frac{\partial}{\partial z}\left(\frac{d}{dt}V_x - KV_y\right) &= -g\frac{\partial \Phi}{\partial x}, \\ \frac{\partial}{\partial z}\left(\frac{d}{dt}V_y + KV_x\right) &= -g\frac{\partial \Phi}{\partial y}. \end{aligned} \tag{7}$$

Equations (I. 7) together with (I. 1) and (I. 3) form a complete set involving only the dependent variables V_x, V_y, V_z, Φ (note that the operator $\frac{d}{dt}$ involves V_x, V_y, V_z: the equations are non-linear). However in the elimination process we have lost a function of integration. If we differentiate the first two of equations (I. 5) with respect to y and x respectively and subtract we obtain, using (I. 2):

$$\begin{aligned} &\left(\frac{\partial}{\partial y} - \frac{\partial \Phi}{\partial y}\right)\left(\frac{d}{dt}V_x - KV_y\right) - \\ &- \left(\frac{\partial}{\partial x} - \frac{\partial \Phi}{\partial x}\right)\left(\frac{d}{dt}V_y + KV_x\right) = 0 \end{aligned} \tag{8}$$

Numerical substitution then shows that in the cases in which we are interested the effects of the operators $\frac{\partial}{\partial x}$, $\frac{\partial}{\partial y}$ swamp those of the multipliers $\frac{\partial \Phi}{\partial x}$, $\frac{\partial \Phi}{\partial y}$ and to a sufficiently close approximation:

$$\frac{\partial}{\partial y}\left(\frac{d}{dt}V_x - KV_y\right) - \frac{\partial}{\partial x}\left(\frac{d}{dt}V_y + KV_x\right) = 0 \tag{9}$$

Elimination of Φ from equations (I. 7) gives simply $\frac{\partial}{\partial z}$ of equation (I. 9) — the approximations in the two cases are consistent. On rearrangement of terms (I. 9) becomes:

$$\begin{aligned} &\mathrm{div}_H \mathbf{v} \cdot (K + \mathrm{curl}_H \mathbf{v}) + \frac{d}{dt}\mathrm{curl}_H \mathbf{v} + \\ &+ \left(\frac{\partial V_z}{\partial x} \cdot \frac{\partial V_y}{\partial z} - \frac{\partial V_z}{\partial y} \cdot \frac{\partial V_x}{\partial z}\right) = 0 \end{aligned} \tag{10}$$

To obtain a fourth equation symmetrical in V_x and V_y we differentiate equations (I. 7) with respect to x and y respectively and add. Then on re-arrangement:

$$\frac{\partial}{\partial z}\left[K\,\mathrm{curl}_H\mathbf{v}-\frac{d}{dt}\,\mathrm{div}_H\mathbf{v}-\left\{\left(\frac{\partial V_x}{\partial x}\right)^2+\right.\right.$$
$$\left.+2\,\frac{\partial V_x}{\partial y}\cdot\frac{\partial V_y}{\partial x}+\left(\frac{\partial V_y}{\partial y}\right)^2\right\}-\left(\frac{\partial V_z}{\partial x}\cdot\frac{\partial V_x}{\partial z}+$$
$$\left.+\frac{\partial V_z}{\partial y}\cdot\frac{\partial V_y}{\partial z}\right)\Big]=g\cdot\nabla_H^2\Phi \qquad (11)$$

where the suffix H in all cases indicates differentiation with respect to x, y only. We shall use equations (I. 1), (I. 3), (I. 10) and (I. 11) as our fundamental set. When other dependent variables are required (e.g., pressure, which may appear in boundary conditions) they are easily computed in terms of our fundamental set.

An important feature of this set of equations is that transformation to a co-ordinate system in uniform horizontal relative motion is almost as simple as in the Newtonian case. V_x and V_y transform as in the latter case by vectorial addition of the relative velocity, V_z and Φ being unchanged. The only difference is that we must add a pressure field (at all levels) whose gradient corresponds to the relative velocity regarded as a geostrophic wind. This is consistent with the assumption that the so-called "tendency equation" is not to be interpreted as an expression of accumulation of mass but rather that the terms associated with this process are usually negligibly small compared with those associated with change of flow — as in classical (subsonic) aerodynamics.

The approximations made could be more convincingly justified by a detailed analysis, but a rigorous proof is possible only *after* obtaining the complete solutions of the approximate equations, when we can check on the precise effect of the omitted terms. A simpler but nevertheless fairly convincing check is the "lifelike" behaviour of the solutions, both qualitatively and quantitatively.

II. Disturbances of Steady Baroclinic Flow

We shall consider first a state of steady baroclinic flow in which the motion is uniform at each level and for simplicity we shall suppose both the "thermal wind" and the static stability constant, i.e., Φ is a linear function of x, y, z. In view of the transformation theorem referred to above there is no loss of generality in supposing:

$$\begin{aligned} V_x &= U(z) : \\ V_y &= V_z = 0 : \\ \Phi &= Ay + Bz : \end{aligned} \qquad (1)$$

So long as we ignore the variability of K the equations of motion are horizontally isotropic — the orientation of the y-axis is irrelevant. Equations (II. 1) are consistent with steady motion if:

$$\frac{dU}{dz}=-\frac{gA}{K} \qquad (2)$$

by reason of (I. 6). The approximation is good except when the thermal wind is very small (or the pressure gradient abnormally large). Since we are interested in the behaviour of strongly baroclinic systems no serious errors are introduced.

We now introduce a small perturbation and write:

$$\begin{aligned} V_x &= U + v_x : \\ V_y &= v_y : \\ V_z &= v_z : \\ \Phi &= Ay + Bz + \varphi : \end{aligned} \qquad (3)$$

where v_x, v_y, v_z, φ are infinitesimal functions of x, y, z, t. Substituting in our fundamental set of equations we obtain the perturbation equations (which are of course linear)

$$\mathrm{div}_H\mathbf{v}+\frac{\partial}{\partial z}v_z=0 :$$

$$K^2(av_y+bv_z)+\frac{d}{dt}(g\varphi)=0 :$$

$$K\,\mathrm{div}_H\mathbf{v}+\frac{d}{dt}\,\mathrm{curl}_H\mathbf{v}+Ka\frac{\partial}{\partial y}v_z=0 :$$

$$\frac{\partial}{\partial z}\left[K\,\mathrm{curl}_H\mathbf{v}-\frac{d}{dt}\,\mathrm{div}_H\mathbf{v}+\right.$$
$$\left.+Ka\frac{\partial v_z}{\partial x}\right]=\nabla_H^2(g\varphi), \qquad (4)$$

where:

$$a \equiv \frac{gA}{K^2}\,;\; b \equiv \frac{gB}{K^2}\,;\; \frac{d}{dt} \equiv U\frac{\partial}{\partial x} + \frac{\partial}{\partial t}\,; \quad (5)$$

By using the identities:

$$\left(\frac{\partial}{\partial z}\frac{d}{dt} - \frac{d}{dt}\frac{\partial}{\partial z}\right) \equiv -Ka\frac{\partial}{\partial x}\,:$$

$$\nabla_H^2 v_y \equiv \frac{\partial}{\partial y}\,\mathrm{div}_H \mathbf{v} + \frac{\partial}{\partial x}\,\mathrm{curl}_H \mathbf{v}\,: \quad (6)$$

we may eliminate successively φ, $\mathrm{curl}_H\mathbf{v}$, $\mathrm{div}_H\mathbf{v}$ from equations (II. 4). Finally we obtain a single partial differential equation involving v_z as the only dependent variable:

$$\frac{d}{dt}\left(K^2 + \frac{d^2}{dt^2}\right)\frac{\partial^2 v_z}{\partial z^2} + 2K^2 a\left(K\frac{\partial}{\partial x} - \right.$$

$$\left. - \frac{d}{dt}\frac{\partial}{\partial y}\right)\frac{\partial v_z}{\partial z} + K^2\left[b\frac{d}{dt}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) - \right.$$

$$\left. - 2Ka^2\frac{\partial^2}{\partial x \partial y}\right] v_z = 0\,: \quad (7)$$

We cannot hope to solve this equation (technically of the fifth order) in complete generality but we may seek certain simple types of solution. Fortunately the simplest solutions are those of greatest practical importance. Since, apart from constants and differential operators (II. 7) involves only functions of z $\left(\text{the function } U(z) \text{ appearing in } \frac{d}{dt}\right)$ it clearly possesses solutions of the form:

$$v_z = N(z)\cdot\Psi\,:\,\Psi \equiv e^{i(\lambda x + \mu y + \vartheta t)}\,; \quad (8)$$

where λ, μ, ϑ are constants and N is a function of z only. In fact we may replace the operators:

$$\frac{\partial}{\partial x} = i\lambda\,;\; \frac{\partial}{\partial y} = i\mu\,;\; \frac{d}{dt} = i(U\lambda + \vartheta)\,; \quad (9)$$

and if at the same time we change our vertical co-ordinate by writing:

$$X \equiv X(z) = \frac{U\lambda + \vartheta}{K}\,;\; \frac{dX}{dz} = -a\lambda\,; \quad (10)$$

we obtain on substitution the ordinary second order differential equation to determine N:

$$X(X^2 - 1)\frac{d^2N}{dX^2} + 2(1 - i\sigma X)\frac{dN}{dX} +$$

$$+ \left[h^2(1 + \sigma^2)X + 2i\sigma\right]N = 0 \quad (11)$$

where we have written:

$$h^2 \equiv \frac{b}{a^2} = \frac{gB}{\left(\frac{dU}{dz}\right)^2}\,:$$

$$\sigma \equiv \frac{\mu}{\lambda}\,: \quad (12)$$

The parameter h^2, which involves both the horizontal and vertical entropy gradients, sums up (apart from matters of scale and boundary conditions) the characteristic properties of the flow. From (I. 4) it is clear that h^2 is simply the Richardson number of the unperturbed flow.

If we can solve (II. 11) with appropriate boundary conditions all the associated perturbation functions are readily determined by means of equations (II. 4). In fact a consistent set of solutions may be obtained in which:

$$v_x = L(z)\cdot\Psi\,:$$

$$v_y = M(z)\cdot\Psi\,:$$

$$\varphi = F(z)\cdot\Psi\,:$$

$$p = G(z)\cdot\Psi, \quad (13)$$

where the pressure perturbation p is determined by:

$$-\frac{1}{\varrho_0}\frac{\partial p}{\partial x} - \frac{d}{dt}v_x - Kv_y - Kav_z \quad (14)$$

or the corresponding equation for $\frac{\partial p}{\partial y}$. Here ϱ_0 is quite uncritical and we may take $\varrho_0 = \varrho_0(z)$ as the mean distribution of density with height in the system with which we are concerned.

For solutions of the form (II. 8) and (II. 13) we obtain the relations:

$$\frac{(1+\sigma^2)}{a} \cdot M = \frac{1}{X}\frac{dN}{dX} + i\sigma\left(\frac{N}{X} - \frac{dN}{dX}\right);$$

$$\frac{(1+\sigma^2)}{a} \cdot L = -\frac{\sigma}{X}\frac{dN}{dX} - i\left(\sigma^2\frac{N}{X} + \frac{dN}{dX}\right);$$

$$\frac{(1+\sigma^2)}{a}\cdot\frac{g}{Ka}\cdot F = \frac{i}{X}\left[\left\{\frac{1}{X}\frac{dN}{dX} + h^2(1+\sigma^2)N\right\} + i\sigma\left\{\frac{N}{X} - \frac{dN}{dX}\right\}\right];$$

$$\frac{(1+\sigma^2)}{a} \cdot \frac{i\lambda}{K\varrho_0} \cdot G =$$

$$= \left(\frac{1}{X}\frac{dN}{dX} + i\sigma\frac{N}{X}\right) + \left(N - X\frac{dN}{dX}\right) \quad (15)$$

With the substitution:

$$N = \left(\frac{1-X}{1+X}\right)^{\frac{i\sigma}{2}} \cdot R \quad (16)$$

equation (II. 11) becomes

$$X(X^2-1)\frac{d^2R}{dX^2} + 2\frac{dR}{dX} + X\left[h^2(1+\sigma^2) + \frac{\sigma^2}{(X^2-1)}\right]R = 0 \quad (17)$$

which is in some ways more convenient.

We will suppose for the moment that λ, μ are real. If at the same time ϑ is real the solution (II. 8) will correspond to a system of stable waves. For the unstable waves we are seeking, ϑ must have a non-vanishing imaginary part and we shall write:

$$\vartheta = \vartheta_0 - i\,\vartheta_1 \quad (18)$$

For such solutions (if any) X defined by (II. 10) becomes a complex variable (with a constant imaginary part) and it is convenient to regard R, determined by (II. 17), as a function of this complex variable. Thus in general L, M, N etc. as well as Ψ are complex numbers. This in no way affects the physical interpretation of v_x, v_y, v_z etc., as the real parts (for example) of the expressions (II. 8) and (II. 13). All it means is that the phase of the wave corresponding to each element (velocity-component, pressure etc.) as well as the amplitude varies with height.

Numerical substitution shows that in normal, convectively stable conditions (we are concerned with mean values over considerable depths) we have:

$$h^2 \gg 1 \quad (19)$$

We shall for the present confine our attention to this, the most interesting case. It can be shown by energy considerations (see below) that in this case we should have:

$$|X^2| \ll 1 \quad (20)$$

over the range for which the perturbations have significant amplitudes. Alternatively we may assume this result and show that our final solutions are consistent with this assumption. Then (II. 17) becomes approximately:

$$\frac{d^2R}{dX^2} - \frac{2}{X}\frac{dR}{dX} - h^2(1+\sigma^2)R = 0 \quad (21)$$

and with the substitutions:

$$H^2 = h^2(1+\sigma^2); \quad Q = HX; \quad (22)$$

we obtain:

$$\frac{d^2R}{dQ^2} - \frac{2}{Q}\frac{dR}{dQ} - R = 0 \quad (23)$$

the general solution of which is:

$$R = a_1R_1 + a_2R_2$$
$$R_1 \equiv e^Q(1-Q)$$
$$R_2 \equiv e^{-Q}(1+Q): \quad (24)$$

where a_1, a_2 are arbitrary constants.

Consider first a hypothetical (but physically possible) system in which motion takes place between two horizontal rigid plane boundaries which, without loss of generality, we may suppose to be at $z = \pm\frac{z_0}{2}$ corresponding to X_1, Q_1, and X_2, Q_2 respectively. We shall suppose the fluid unbounded in any horizontal

direction. Then for the functions to be finite at infinity we must take λ, μ real. At $z = \pm \frac{z_0}{2}$ the normal velocity v_z must vanish which will be the case if:

$$R = 0;\ Q = Q_1,\ Q_2 \qquad (25)$$

Then we have:

$$-\frac{a_2}{a_1} = e^{2Q_1}\left(\frac{1 - Q_1}{1 + Q_1}\right) = e^{2Q_2}\left(\frac{1 - Q_2}{1 + Q_2}\right): \qquad (26)$$

and if we write:

$$(Q_2 - Q_1) = 2\alpha:$$
$$(Q_2 + Q_1) = -2i\beta, \qquad (27)$$

we have:

$$\beta^2 = 2\alpha \coth 2\alpha - 1 - \alpha^2 \equiv$$
$$\equiv (\alpha - \tanh \alpha)(\coth \alpha - \alpha) \qquad (28)$$

as the condition to be satisfied if solutions are to exist. On substitution for Q_1, Q_2 in (II. 27) we have:

$$2\alpha = \frac{H}{K} \cdot \frac{dU}{dz}\lambda z_0:$$
$$-2i\beta = \frac{H}{K}[\lambda(U_1 + U_2) + 2\vartheta_0]$$
$$-2i\vartheta_1 \cdot \frac{H}{K} \qquad (29)$$

so that α is necessarily real. It follows from (II. 28) that β is either purely real or purely imaginary. Only when β is real, i.e. when

$$|\alpha| < \alpha_0 = 1.1997\ [\alpha_0 = \coth \alpha_0] \qquad (30)$$

do unstable solutions exist. Then by (II. 29):

$$\vartheta_1 = \beta \cdot \frac{K}{H}: -\frac{\vartheta_0}{\lambda} = \frac{(U_1 + U_2)}{2} \qquad (31)$$

where U_1, U_2 are the unperturbed velocities at levels $z = \pm \frac{z_0}{2}$ respectively. The second of these equations shows that the waves travel with the mean unperturbed current (we may call the level $z = 0$ the "steering level"). The first equation determines the growth-rate. It is easily shown that $|\beta|$ defined by (II. 28) has a maximum for a particular value of α. Then:

$$|\beta| = 0.3098$$
$$|\alpha| = 0.8031 \qquad (32)$$

For these values, with the additional condition $\sigma = 0$, the growth-rate is a maximum. At the same time (II. 29) determines the wavelength $\frac{2\pi}{\lambda}$ in terms of the parameters of the unperturbed system. The disturbances are easily seen to be, at each level, a series of (growing) ridges and troughs with their axes at right angles to the unperturbed thermal wind. Since the structure of these disturbances is very similar to that of the disturbances of more realistic systems it is of interest to examine them in detail. This structure is most conveniently described by graphs showing the variation with height of the phase and amplitude of the waves representing various salient features such as the pressure perturbation, vertical velocity, etc. The graphs for p, φ, v_z, $\mathrm{div}_H \mathbf{v}$ are shown in Fig. 1. The distribu-

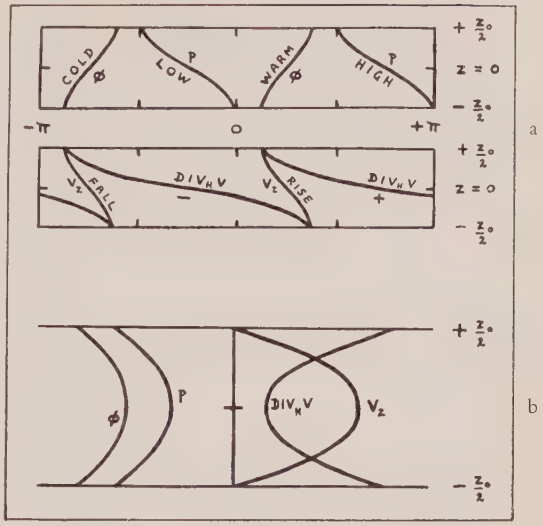


Fig. 1 a. Phase Variations: Above p, φ; Below v_z, $\mathrm{div}_H \mathbf{v}$;
Fig. 1 b. Amplitude Variations: Left p, φ; Right v_z, $\mathrm{div}_H \mathbf{v}$.

tion for v_y may be inferred directly from the pressure field since it is easily shown that for these disturbances [and in fact for the disturbances of all systems satisfying (II. 19)] the winds are to a first approximation geostrophic. Ageostrophic winds (including in the present instance v_z) are of order $\frac{1}{h}$.

It will be observed that the pressure trough slopes upwards and backwards in the atmosphere while the warm tongue (entropy maximum) slopes upwards and forwards. At low levels the warm tongue is slightly ahead of the pressure trough but at high levels the warm tongue is slightly to the rear of the upper pressure ridge. Upward motion (and hence, potentially, rainfall) is a maximum at middle levels $^1/_8$ wavelength ahead of the surface pressure trough. We may note that at middle levels v_y and v_z are in phase, the combined motion being upwards towards cold, downwards towards warm air. We may note also that at the same level v_z and φ are in phase with rising warm and descending cold air, corresponding to a decrease in potential energy. In fact on integration over a whole wavelength we find that there is a positive correlation between v_z and φ and a net decrease in the potential energy of the system as a result of the disturbance. It is of course this release of potential energy which feeds the kinetic energy of the growing disturbance. Our analysis shows that such a process (similar to that conceived by Margules) is consistent with all the equations and constraints of motion and in fact that such processes must occur from time to time.

It is easily verified by substituting typical values of the parameters that if z_0 is taken as the height of the tropopause the wavelength of the disturbance of maximum growth-rate is of approximately the same size as observed long waves. For smaller values of the static stability (as in large cloud masses) and smaller vertical extents we obtain "dominant" wavelengths of the order of magnitude of observed extratropical wave-cyclones. Thus we are certainly concerned with disturbances of the right order of magnitude. Our systems are not yet however sufficiently realistic for positive identification. As a first step towards realism we remove the artificial boundary at $z = +\frac{z_0}{2}$ and consider a system in which the atmosphere extends upwards indefinitely but at some definite level the static stability increases abruptly. For mathematical simplicity we take the "thermal wind" to be the same in both the lower and upper "regimes". This system is unsymmetrical so we shall put $z = 0$ at the earth's surface (rigid boundary) and $z = z_0$ at the boundary between the two regimes. Thus $z = z_0$ might correspond to an inversion or stabilisation of lapse-rate — were it not for the observed thermal wind change z_0 might be the height of the tropopause. Alternatively, the lower regime might be a (baroclinic) cloud mass of low base (small effective static stability) surmounted by unsaturated air. In all these cases the Richardson number (h_1^2) in the lower "regime" is less than that (h_2^2) in the upper regime. We can write down, as before, the general solution of (II. 21) appropriate to each regime. Clearly for continuity at the boundary we must have λ, μ, ϑ (and therefore X but not Q) the same in each regime. We still have two more arbitrary constants than before, but we have two additional "internal" boundary conditions since both the phase and amplitude (one complex number) of p and v_z (normal velocity) must be continuous at $z = z_0$. Our upper boundary condition is now $v_z \to 0$ as $z \to \infty$ for it is clear from (II. 24) that one of R_1, R_2 increases and the other decreases exponentially with height. (The boundary condition ensures that *all* the perturbation functions decrease exponentially.) Finally, we obtain, as before, a relation between λ and ϑ. If we write:

$$2\alpha = h_1 \sqrt{1 + \sigma^2}\, \frac{\lambda}{K}\, \frac{dU}{dz}\, z_0,$$

$$\beta = h_1 \sqrt{1 + \sigma^2} \cdot \frac{\vartheta_1}{K},$$

$$k_1 = \frac{h_1}{h_2}, \qquad (33)$$

then

$$\beta^2 = \frac{(1 - k_1^2)\,(2\alpha - \tanh 2\alpha)}{(k_1 + \tanh 2\alpha)} - \left(\alpha - \frac{k_1}{2}\right)^2 \qquad (34)$$

replaces (II.28) while for the "steering level" as defined above:

$$z_\omega = \frac{z_0}{2}\left(1 + \frac{k_1}{2\alpha}\right) \tag{35}$$

If we put $h_2 = \infty$ (infinite static stability) equations (II.34) and (II.35) reduce to formulae appropriate to the "two rigid boundary" system and the disturbance is confined to the lower regime. As k_1 increases from zero the disturbance gradually extends into the upper regime but as there is always exponential decrease with height in the upper regime, provided $k_1 < 1$, then, except near this limit, the actual conditions much above $z = z_0$ are quite uncritical. In many practical cases k_1 is nearer zero than unity and then the disturbance in the lower regime does not differ greatly from that of our original system (the limiting case). For a given value of k_1 we find as before that β is real only for sufficiently long waves (α sufficiently small) and for one particular wavelength β (and therefore growth-rate) is a maximum. As k_1 increases from zero β decreases slowly at first from its limiting value but vanishes when $k_1 = 1$. For $k_1 > 1$ there are no unstable solutions. (It would appear to be a general result that for instability of this type the Richardson number must have a minimum value within a certain finite region; the resulting disturbances then have their maximum amplitude in this region with exponential decrease in the surrounding regions. These conditions are satisfied in practice as a general rule.)

Similar calculations may be made for systems of three (or more) horizontally stratified regimes. An easily investigated system, representing the opposite extreme to the two-regime system discussed above, is that in which the static stability is relatively small within a layer which we take to be between levels $z = \pm\frac{z_0}{2}$ and relatively large both above and below. The outer regimes we take (for simplicity) to be of indefinite extent. This system corresponds, for example, to a baroclinic cloud mass of very high base. For simplicity we take $h = h_2$ in each of the outer regimes, $h = h_1$ in the inner regime. Then with the same notation as before:

$$\beta^2 = \frac{(1-k_1^2)\{2\alpha(\coth 2\alpha + k_1) - (1-k_1^2)\}}{(1+k_1^2+2k_1\coth 2\alpha)} - \alpha^2 \tag{36}$$

The general behaviour is similar to that of the two-regime system except that now the disturbance decreases exponentially in both directions away from the inner regime. (For small values of k_1 it is almost as if the conditions at the boundaries of the inner regime were independent.) Thus disturbances developing on a high level cloud sheet, for example, would be unnoticed at ground level in their early stages.

Table 1 gives values of α, β corresponding to the "dominant" disturbance of maximum growth-rate calculated for $k_1^2 = 1/6$, typical of the change from cloud to unsaturated air, for the three systems discussed above. Note the relatively small decrease in β due to "losses" at "imperfectly rigid" boundaries.

Table 1

System	I	II	III
α	0.8031	0.6190	0.4445
β	0.3098	0.2988	0.2885

The structure of the disturbances of the more complex systems may be studied in the same way as before. Figs. 2 and 3, to be interpreted as Fig. 1, correspond to the two-regime and three-regime systems respectively.

An interesting feature of the two-regime system appears when we compute the displacement of the internal boundary (the base of the change in lapse-rate) at $z = z_0$ due to the growing disturbance. This boundary (which might correspond roughly to the tropopause) is sucked down in the vicinity of the upper (i.e., at $z = z_0$) pressure minimum and pushed up in the vicinity of the upper pressure maximum (the phase coincidence becomes exact as $k_1 \to 0$, otherwise it is a close approximation). This result (opposite to that to be expected if cyclonic vorticity were generated by a simple vertical "stretching" — but note that behaviour in the upper regime, e.g., the stratosphere, *is* consistent with this view) is in good agreement with observed behaviour.

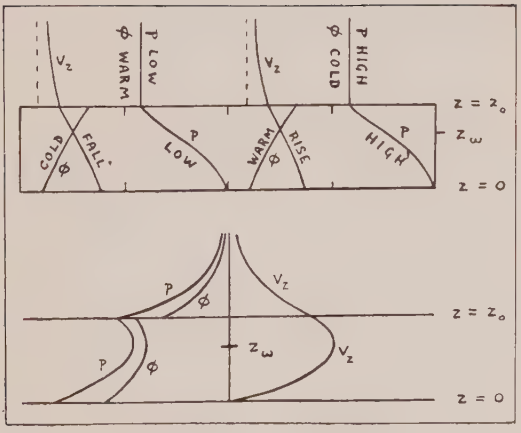


Fig. 2. Above: Phases. Below: Amplitudes.

In the above we have considered only the case of uniform thermal wind, the changes in Richardson number being due to changes in static stability. The general case, when both thermal wind and static stability change, is much less simple mathematically though it could probably be tackled by numerical methods. But from energy considerations there seems little doubt that the general behaviour of any system depends primarily on the distribution of Richardson number and much less on the way in which it is compounded.

We have still only considered systems which are horizontally of infinite extent and although strongly baroclinic regimes are often of considerable longitudinal extent (in the direction of the thermal wind), they are seldom very broad. As a further step towards realism we consider a three-regime system in which the regimes are now side by side. For simplicity we commence with the case of motion between horizontal rigid boundaries which we found previously to be a useful first approximation. We suppose the inner regime to occupy $-\frac{y_0}{2} < y < \frac{y_0}{2}$ and the outer regimes to be of indefinite extent. All three regimes occupy $-\frac{z_0}{2} < z < \frac{z_0}{2}$. Once again we suppose differences in Richardson number to be due solely to differences in static stability (consistent with a continuous temperature distribution if the stability reduction is due to cloud). We take $h = h_1$ in the inner regime, $h = h_2$ in each outer regime with $|h_1| < |h_2|$ so that the inner regime corresponds, for example, to a baroclinic cloud mass. In previous cases we found the only restriction on σ was that it must be real, though the most interesting case (maximum growth-rate) was $\sigma = 0$. In the present instance we find simple solutions only if:

$$h_1^2 (1 + \sigma_1^2) = h_2^2 (1 + \sigma_2^2) \equiv H^2 \qquad (37)$$

and then $Q = HX$ has the same interpretation everywhere. The boundary conditions at $z = \pm \frac{z_0}{2}$ are satisfied as in our first problem and then they are satisfied in all regimes simultaneously. The internal boundary conditions require continuity of p and v_y (normal velocity) at $y = \pm \frac{y_0}{2}$. (Since the winds are roughly geostrophic these conditions are nearly equivalent — hence the "fit" must be

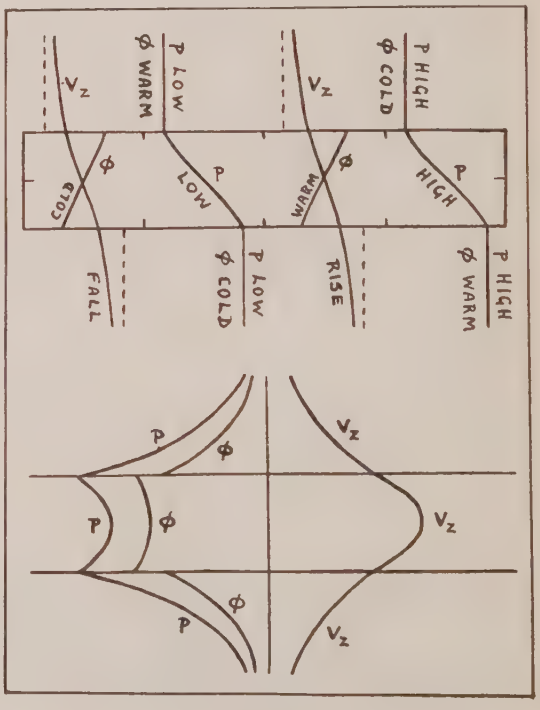


Fig. 3. Above: Phases. Below: Amplitudes.

correct to a higher order in $\frac{1}{h}$). We shall suppose (and our solutions require) that σ_1 is real, σ_2 purely imaginary, consistent with our assumption that $k_2 \equiv \frac{h_1}{h_2} < 1$. Then the boundary conditions at $y = \pm \infty$ are satisfied (all perturbations decreasing exponentially) with appropriate (opposite) choice of roots $\pm \sigma_2$ in the two outer regimes. In the inner regime we take a linear combination of the solutions corresponding to $\pm \sigma_1$ respectively and then have sufficient arbitrary constants to satisfy all the boundary conditions if, at the same time:

$$\frac{i\sigma_2}{\sigma_1} = \tan\left(\frac{\lambda\sigma_1}{2}\gamma_0\right) \qquad (38)$$

or, by (II. 37)

$$\sigma_1^2\left[k_2^2 + \tan^2\left(\frac{\lambda\gamma_0}{2}\sigma_1\right)\right] = 1 - k_2^2 \qquad (39)$$

Since $k_2^2 < 1$ this equation always possesses at least one real root for $|\sigma_1|$. (If there is more than one we take the smallest — corresponding to maximum growth-rate.) For the growth-rate we have:

$$\vartheta_1 = \beta' \frac{K}{h_2}:$$

$$\beta' \equiv \frac{1}{k_2} \cdot \frac{\beta}{\sqrt{1+\sigma_1^2}}, \qquad (40)$$

where β is defined by (II. 28). Once again β' is real, and the disturbance unstable, only within the range (II. 30) and for one particular value of α growth-rate is a maximum. Since σ_1 depends on λ the "dominant" values of $|\alpha|$ and $|\beta|$ are slightly different from those given by (II. 32) — $|\alpha|$ is somewhat greater and $|\beta|$ slightly smaller by an amount depending on $\frac{\gamma_0}{z_0}$ and on k_2. The variation of β' with α as compared with the case of an infinite cloud-sheet ($\gamma_0 = \infty$) is shown in Fig. 4 for a typical case in which $k_2^2 = \frac{1}{10}$; $\frac{1}{\sqrt{b_2}}\frac{\gamma_0}{z_0} = \frac{1}{4}$; (see II. 5).

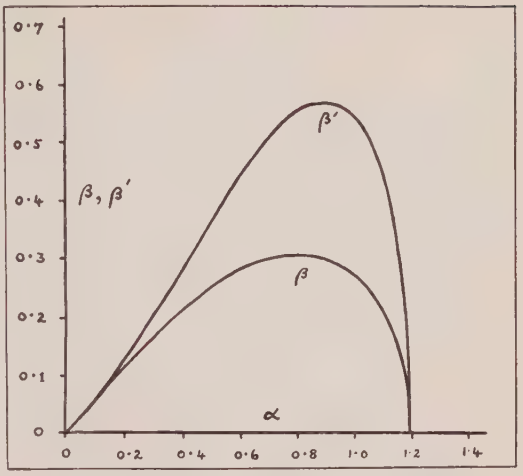


Fig. 4. Selection Curves.

It will be observed that the "selectivity" (i.e. the sharpness of the maximum) is increased. As a numerical example we may compute the characteristic features of the dominant wave of the system to which Fig. 4 applies, assuming in addition $K = 0.4$ hr^{-1} (lat. 50° approximately) : $\sqrt{b_2} = 115$ (typical for unsaturated air) : $z_0 = 5$ km : $\gamma_0 = 150$ km (height and width of cloud mass) : $\frac{dU}{dz} = 10$ hr^{-1}. Then we derive:

$$h_2^2 = 21;\ \alpha = 0.893;\ \beta' = 0.575;\ \sigma_1 = 1.333; \qquad (41)$$

and then:

$$\frac{1}{\vartheta_1} = 19.9 \text{ hr};\ \mathbf{L} \equiv \frac{2\pi}{\lambda} = 1{,}070 \text{ Km}, \qquad (42)$$

so that the disturbance doubles its size in approximately 14 hours. The growth-rate is slightly less than that sometimes observed since in practice the effective Richardson number may be smaller than that assumed (and then our approximations are not so good — but see below). But both growth-rate and wavelength are of the right order of magnitude for cyclone waves. On the other hand when z_0 is the height of the tropopause and the Richardson number is not much less than that appropriate to unsaturated air (only cloud masses comparable in size with the disturbance pro-

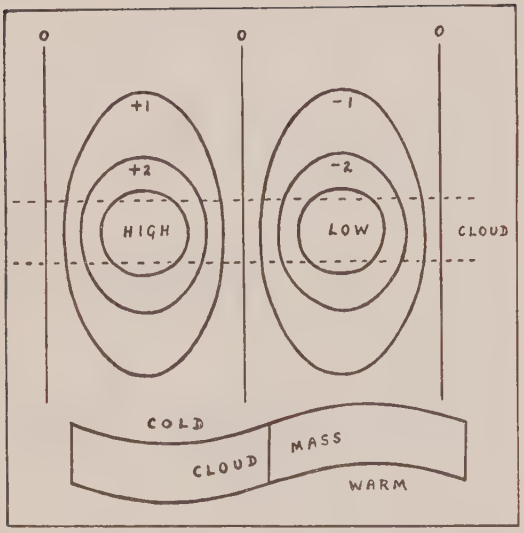


Fig. 5. Pressure-field and Cloud-mass Perturbation.

duce a large reduction in effective Richardson number) we find that **L** is more than doubled (in a typical case we find **L** is of the order of 4,000 Km) while ϑ_1 is considerably reduced $\left(\text{in a typical case } \frac{1}{\vartheta_1} = 60 \text{ hr}\right)$. These values are in good agreement with observed values in long waves. In general we find that growth-rate is intermediate between the values appropriate to infinite sheets of Richardson number h_1 and h_2 respectively, approaching the former for very broad, the latter for very narrow "cloud belts".

The structure of the "dominant" disturbance has many interesting features which will be discussed only briefly. The vertical distribution of phase and amplitude of the perturbation functions is exactly as in Fig. 1, but the horizontal structure is of course more complex. The approximate pressure-perturbation pattern $\left(\text{correct to zero order in } \frac{1}{h}\right)$ is shown in Fig. 5.

Apart from the changes of phase and amplitude indicated in Fig. 1, the pattern is the same at all levels. For comparison with observation we must superpose the unperturbed pressure field. We then find that (except when $U = 0$) closed centres are absent until the disturbance has attained a definite size (different at different levels and perhaps never attained, especially at high levels, — but this question cannot be discussed on a theory of small perturbations, which applies only to the early stages of development). Since the winds are roughly geostrophic $\left(\text{correct to zero order in } \frac{1}{h}\right)$ we may infer the approximate horizontal wind field from Fig. 5. But note that the trajectories are not even approximately along the isobars except at the steering level $z = 0$. Relative to the disturbance the air is blowing through from, say, the east at low levels and the west at high levels. The combined relative motion is associated with a deformation of the cloud mass. This deformation is shown below Fig. 5 (since it is "infinitesimal"), in correct phase for the level $z = -\frac{z_0}{2}$ (corresponding to the earth's surface). The phase varies with height and is the same as that of the entropy perturbation (warm tongue) at levels $z = 0, \pm\frac{z_0}{2}$, the phase-difference elsewhere being very small. Clearly the cloud-mass corresponds to the frontal region of the growing cyclone and its displacement towards cold air to the boundary of the "warm sector". It should be emphasised that we are concerned here only with the *broad* features of disturbances. Errors of detail are inevitable since actual initial systems are usually more complicated in structure than we have assumed. From this point of view it does not seem to matter very much whether or not the velocity field contains discontinuities (except when these are unusually large *and* extensive) provided the smoothed fields are the same. Moreover sharp discontinuities observed in practice are often the result of rather than the prerequisite for development. When the theoretical fields are more accurately computed (correct to the first order in $\frac{1}{h}$, the maximum attainable with present approximations) we find discontinuities of wind, pressure gradient, entropy developing along each surface of the cloud mass. This behaviour may be regarded as the realisation of a latent discontinuity — in effective static stability — between saturated

and unsaturated air. These discontinuities are more complex in structure (and in some ways in better agreement with observation) than those at the plane surfaces ascribed to theoretical "fronts". Moreover the genesis of discontinuity by development is consistent with the observed "sharpening" of fronts during cyclogenesis. But a satisfactory theory of frontogenesis cannot be based on considerations of small disturbances and we must leave this aspect (from our point of view a matter of detailed structure) for the present.

It is easy to combine the virtues (from the point of view of realism) of the horizontally and vertically "stratified" systems by considering the systems shown in cross-section in Fig. 6. The only difference as compared with the system last considered, is that in place of Fig. 1, we use Figs. 2 and 3 respectively for the vertical variations of phase and amplitude provided that the conditions:

$$\frac{h_2^2}{h_1^2} = \frac{h_4^2}{h_3^2} \tag{43}$$

are satisfied. We then find that all the boundary conditions can be satisfied simultaneously. The condition (II. 43) is necessary for solutions of simple mathematical form but it has little physical significance since very little perturbation-energy is associated with the "corner" regimes. In fact when h_1^2 is considerably smaller than the value of h^2 outside it is this alone (together with the dimensions of the cloud mass) which is the main determining factor of the features of the disturbance.

Several refinements and extensions of the theory will not be discussed in detail here. Thus the solutions of (II. 17) correct to order $\frac{1}{h^3}$ are easily determined. Applied to the first system considered they yield more precise formulae for growth-rate etc. It appears that serious errors do not result until h^2 approaches fairly close to unity. For $h^2 < 1$ a second type of instability (corresponding to "vertical overturning" — see below) becomes possible. This type of instability is associated with development on a smaller scale (motion in the vicinity of cold fronts, tropical cyclones etc.) and will be discussed elsewhere. Another extension is the calculation of second-order

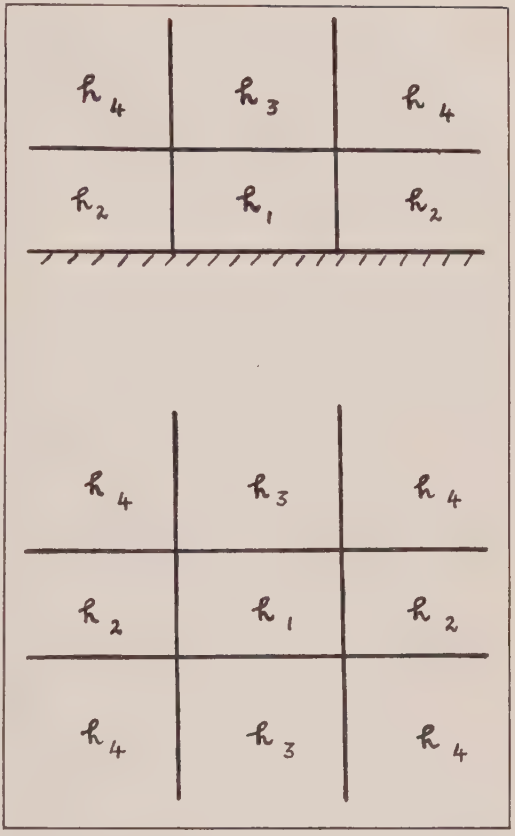


Fig. 6. 6- and 9-regime Systems (Vertical Section).

perturbations i.e. correct to the second order of small quantities. We then obtain terms involving Ψ^2 i.e., second harmonic terms. Calculations are simple for the first system, corresponding to Fig. 1. We find that the complete perturbation is no longer symmetrical as between "high" and "low". At all levels the pressure troughs are accentuated and the ridges flattened, in good agreement with observed behaviour. In fact, correct to zero order in $\frac{1}{h}$, the phase-lines of the first and second order perturbations of pressure coincide (for troughs) at all levels. We also obtain terms independent of Ψ, corresponding to "transport" phenomena. Thus for example the disturbances transport heat *upwards* and their final effect must be to *increase* the static stability of the system. (In this case of course

the same result is obtained from the integrated correlation of v_z and φ.) An important corollary is that since large-scale disturbances are always transporting heat upwards we must, for statistical balance, have net cooling of the upper troposphere by radiation. Independent calculations of radiation flux have led to the same conclusion.

III. Long Wave Modifications

In the previous analysis, in which we have neglected the variability of the Coriolis parameter, we found that the disturbances move with the unperturbed current at the "steering level" which, in the case of symmetrical systems (Figs. 1 and 3) is the middle level of the system. In the system to which Fig. 2 applies the steering level is somewhat elevated, corresponding to a limited extension of the disturbance into the upper layer. If we apply the superposition (of uniform wind and corresponding geostrophic pressure gradient) theorem referred to in the first section we obtain the law of "contour steering" (*not*, as is often stated "thermal steering") at the steering level. So far at least as direction of travel is concerned this is in good agreement with observation (though we have proved the result only for nascent disturbances). But this law applies only to disturbances on not too large a scale. As we shall see, the most important new feature arising when we take into account the variation of K with latitude is a modification in the steering law.

For dealing with the more complicated problems in which we take into account the variability of the Coriolis parameter, or use a polar co-ordinate system, or are dealing with complicated boundary conditions, etc., it is convenient to reformulate the perturbation equations. This involves further approximations so that the solutions are valid only to zero order in $\frac{1}{h}$ (on the other hand in long wave problems h^2 is much larger than in cyclone problems) and the calculations are not so easily extended to the second order of small quantities but the differential equations are of a simpler form, are better adapted to numerical work and the solutions still retain the essential features we are seeking to determine. We commence with a "standard" distribution $\varrho_0(z)$, which may be the actual distribution in some central part of the region in which we are interested, and we shall suppose $p_0(z)$ the corresponding pressure distribution. Let p', ϱ' be the differences at any point from the "standard" values at the same level. Then $\frac{1}{\gamma} \cdot \frac{p'}{p_0}$, $\frac{\varrho'}{\varrho_0}$ are both small compared with unity and normally in regions of strong thermal wind the former is much smaller than the latter. Hence approximately:

$$\varphi = \frac{1}{g\varrho_0} \cdot \frac{\partial p}{\partial z} = \frac{1}{g} \frac{\partial}{\partial z}\left(\frac{p}{\varrho_0}\right) \qquad (1)$$

for perturbations of sufficiently shallow systems, in which $\left|\frac{\partial}{\partial z} \log p\right| \gg \left|\frac{d}{dz} \log \varrho_0\right|$. If in the initial system the static stability and thermal wind are uniform we have for adiabatic motion (cf. II. 4):

$$a v_y + b v_z + \frac{1}{K^2} \cdot \frac{d}{dt} \frac{\partial}{\partial z}\left(\frac{p}{\varrho_0}\right) = 0 \qquad (2)$$

We have noted that the perturbation winds are, to zero order in $\frac{1}{h}$, geostrophic. Then replacing v_y by $\frac{1}{K\varrho_0} \cdot \frac{\partial p}{\partial x}$ and differentiating:

$$b \frac{\partial v_z}{\partial z} + \frac{1}{K^2}\left(\frac{\partial}{\partial z} \frac{d}{dt} + Ka \frac{\partial}{\partial x}\right) \frac{\partial}{\partial z}\left(\frac{p}{\varrho_0}\right) = 0 \qquad (3)$$

neglecting once again the variation of $\log \varrho_0$. Hence by the first of equations (II.4) and (II.6):

$$b \cdot \mathrm{div}_H \mathbf{v} = \frac{1}{K^2} \cdot \frac{d}{dt} \frac{\partial^2}{\partial z^2}\left(\frac{p}{\varrho_0}\right) \qquad (4)$$

When h^2 is large the third of equations (II.4) approximates to:

$$-\mathrm{div}_H \mathbf{v} = \frac{1}{K} \cdot \frac{d}{dt} \mathrm{curl}_H \mathbf{v} = \frac{1}{K^2} \cdot \frac{d}{dt} \nabla_H^2 \left(\frac{p}{\varrho_0}\right) \qquad (5)$$

on substitution of the geostrophic winds. Combining (III. 4) and (III. 5)

$$\frac{d}{dt}\left[\frac{\partial^2 p}{\partial x^2}+\frac{\partial^2 p}{\partial y^2}+\frac{1}{b}\frac{\partial^2 p}{\partial z^2}\right]=0 \qquad (6)$$

and this condition is certainly satisfied if:

$$\frac{\partial^2 p}{\partial x^2}+\frac{\partial^2 p}{\partial y^2}+\frac{\partial^2 p}{\partial(\sqrt{b}\cdot z)^2}=0 \qquad (7)$$

All the other perturbation functions are readily expressible in terms of p and using (III. 7) in place of (II. 7) we can obtain, correct only to zero order in $\frac{1}{h}$, the results of section II.

Our present approximations are therefore consistent with, albeit more drastic than those made earlier. Now if we take into account the variability of K we have, in place of (III. 5):

$$-\operatorname{div}_H \mathbf{v}=\frac{1}{K}\frac{d}{dt}(K+\operatorname{curl}_H \mathbf{v})=$$

$$=\frac{1}{K^2}\left[\frac{d}{dt}\nabla_H^2\left(\frac{p}{\varrho_0}\right)+\frac{dK}{dy}\cdot\frac{\partial}{\partial x}\left(\frac{p}{\varrho_0}\right)\right] \qquad (8)$$

and then, in place of (II. 6)

$$\frac{d}{dt}\left\{\nabla_H^2 p+\frac{1}{b}\frac{\partial^2 p}{\partial z^2}\right\}+\frac{dK}{dy}\cdot\frac{\partial p}{\partial x}=0 \qquad (9)$$

For solutions of the type (II. 8 and II. 13) studied in the previous section (III. 9) leads to the ordinary differential equation:

$$\frac{d^2G}{dP^2}=\left(\frac{1}{C^2}-\frac{1}{P}\right)G \qquad (10)$$

where

$$C^2\equiv\frac{h^2}{(\lambda^2+\mu^2)}\cdot\left(\frac{1}{K}\frac{dK}{dy}\right)^2$$

$$P\equiv CQ=Ch\sqrt{1+\sigma^2}\cdot X. \qquad (11)$$

From (III. 2) and (III. 10) we obtain the relation:

$$N=-P^2\frac{d}{dP}\left(\frac{G}{P}\right) \qquad (12)$$

and in place of (III. 10) we may use:

$$Q(Q-C)\frac{d^2N}{dQ^2}-(2Q-C)\frac{dN}{dQ}-$$

$$-(Q-C)^2N=0 \qquad (13)$$

which of course reduces to (II. 23) when $C=0$. (To our present approximation N and R are equivalent.) The general solutions of (III. 10) and (III. 13) are expressible in terms of Whittaker functions but since we are interested in functions of a complex variable evaluation of the solutions and determination of the dominant solution is in general a laborious process. If however, in order to discover the initial effect of the term involving $\frac{dK}{dy}$ when the "correction" is not too large, i.e. for waves which are not too long, we assume $|C|\ll 1$ then (III. 13) approximates to:

$$Q\frac{d^2N}{dQ^2}-2\frac{dN}{dQ}-QN=$$

$$=C\left[\frac{d^2N_0}{dQ^2}-\frac{1}{Q}\frac{dN_0}{dQ}-2N_0\right] \qquad (14)$$

(where N_0 corresponds to $C=0$), which is easily solved by variation of parameters (the L. H. S. is the same as in II. 23), the solutions involving exponential integrals. If we consider now the first initial system (two rigid boundaries) and compare our solutions for C small but non-vanishing with our original solutions for $C=0$ we find that the dominant wavelength and the corresponding value of ϑ_1 are unaltered, but that we obtain an additional term in the real part of the wave-velocity, corresponding to a lowering of the steering-level, given by:

$$-\delta U_\omega=0.726\cdot\frac{1}{(\lambda^2+\mu^2)}\cdot\frac{dK}{dy} \qquad (15)$$

a formula differing only by a numerical factor from that applicable to (hypothetical) barotropic waves. The formula appears to be in reasonably good agreement with observation in middle-high latitudes (poleward of 45° lat.) where the dominant wavelength is typically about 4,000 Km (though we have proved the result only for *growing* waves).

The steering-level corresponds roughly with minimum amplitude of the pressure perturbation (cf. Fig. 2) and the present lowering of the steering-level is associated with an increase of perturbation amplitude at high levels (i.e., towards the tropopause) and a decrease at low levels. That long waves are more intense at high as compared with low levels is of course well known.

We have observed that the decrease of mean density with height involves a "distortion" which is most serious in the case of deep waves such as long waves. An approximation to the modifications involved is obtained by replacing the first of equations (II. 4) by:

$$\mathrm{div}_H \mathbf{v} + \frac{\partial v_z}{\partial z} - \frac{1}{z_c} \cdot v_z = 0 \qquad (16)$$

where

$$\frac{1}{z_c} = -\frac{d}{dz} \log \varrho_0 = \frac{g}{\gamma R T} \qquad (17)$$

Then in place of (II. 23) we obtain finally:

$$\left[Q \frac{d^2 R}{dQ^2} - 2 \frac{dR}{dQ} - QR \right] + \\ + m \left[Q \frac{dR}{dQ} - 2 R \right] = 0 \qquad (18)$$

where

$$m \equiv \frac{1}{Ha \lambda z_c} = -\frac{1}{2 \alpha} \cdot \frac{z_0}{z_c} \qquad (19)$$

(cf. II. 29, II. 5 and II. 2). Once again we shall consider the initial form of the "correction" when it is small, i.e. when z_0 is appreciably less than 1.6 z_c. Making the substitution:

$$R = e^{\frac{m}{2} Q} \cdot R' \qquad (20)$$

and neglecting m^2 we get in place of (III. 18):

$$Q \frac{d^2 R'}{dQ^2} - 2 \frac{dR'}{dQ} - (Q + m) R' = 0 \qquad (21)$$

Using once again the method of variation of parameters we obtain the general solution of (III. 21) and can study the modifications resulting from small but non-vanishing m as compared with our original solutions when $m = 0$. As above, when we were concerned with the effect of $\frac{dK}{dy}$, we find that the formulae for dominant wavelength and growth-rate are unaffected but there is an additional term in the real part of the wave-velocity. Expressed in terms of the steering-level we find that the latter is depressed by an amount $-\delta z$ where:

$$-\frac{\delta z}{z_0} = 0.0875 \left(\frac{z_0}{z_c} \right) \qquad (22)$$

Even when, as is the case in practice, $\frac{z_0}{z_c}$ is comparable with unity, and the correction is rather rough, this depression of the steering-level due to decrease of density with height is much smaller, for long waves, than that due to variation of the Coriolis parameter. To our present approximations these effects are of course additive. We may note that the factor $e^{-\frac{m}{2} Q}$ in (III. 20) is simply $\frac{1}{\sqrt{\varrho_0}}$ multiplied by a (complex) constant. Thus apart from the difference in behaviour of the amplitudes of perturbations which are functions of R' as compared with those of the R of (II. 23), a relatively minor difference, we find that perturbation amplitudes are multiplied by $\frac{1}{\sqrt{\varrho_0}}$, corresponding to increase in relative amplitude with height (and, to this extent, constancy of wave-energy density) and further accentuating the feature, already noted above, that long wave amplitude increases (on the whole) with height. Of course the corrections (III. 15) and (III. 22) apply to cyclone waves as well as long waves. But since both corrections are proportional to the square of the wavelength they can usually be neglected for practical purposes in the former case.

For practical application we have identified z_0 for long waves as the height of the tropopause since the mean value of h^2 over a considerable depth of the stratosphere is normally considerably larger, in significant regions, than in the troposphere. For a closer approximation we may study a system, similar to that to which Fig. 2 applies, in which the disturbance extends (but with exponential

decrease with height) into the stratosphere. We cannot immediately apply our previous results since the sign of $\frac{dU}{dz}$ in the upper regime is reversed and the internal boundary (i.e., the tropopause) is no longer horizontal so that a precise mathematical formulation leads to a more complicated problem. Nevertheless it is easy to see that the two cases are not very different (and incidentally provide confirmatory evidence that the essential characteristics of a system depend on the Richardson number rather than its component elements). For the solutions R_1, R_2 (equation II. 24) are simply interchanged (h replaced by $-h$) when $\frac{dU}{dz}$ is reversed. In order to satisfy the boundary conditions at $z = \infty$ we have to choose the same solution (i.e., that which decreases with height) as before. We shall suppose the change in h^2 at the tropopause large (k_1 small) so that decrease in the stratosphere is rapid and most of the wave-energy in this regime concentrated just above the tropopause. The distribution of $\text{curl}_H \mathbf{v}$ in the lower regime, and in particular at the tropopause, cannot differ much from the limiting case (Fig. 1 — cf. Fig. 2). Since there are no discontinuities of velocity in our initial system we find that $\text{curl}_H \mathbf{v}$ must be continuous (at least to zero order in $\frac{1}{h}$, the winds being roughly geostrophic) at the tropopause. Hence $\text{curl}_H \mathbf{v}$ in the lower stratosphere is not much altered when $\frac{dU}{dz}$ is reversed. And just above the tropopause U is only slightly reduced by this reversal. Hence $\frac{\partial v_z}{\partial z}$, which is determined to this order of accuracy, by $\left(U\frac{\partial}{\partial x} + \frac{\partial}{\partial t}\right)$ $\text{curl}_H \mathbf{v}$ is not much altered and this is consistent with our choice of R_1 or R_2 corresponding to v_z decreasing. In Fig. 2 both v_y and v_z are continuous at the internal boundary, the resultant velocity having a slope smaller than that of the isentropic surfaces in the lower regime but greater than that in the upper regime and producing the 180° phase change in φ. When $\frac{dU}{dz}$ in the stratosphere is reversed the slope of the isentropic surfaces there is reversed but there is no qualitative (and only a relatively small quantitative) change in the above description. In fact in the significant region (i.e., just above the tropopause) we can reverse the sign of $\frac{dU}{dz}$ and our original solution (Fig. 2) is still a rough approximation to the required solution. Alternatively we may imagine the stratospheric value of $\frac{dU}{dz}$ changing gradually through zero. The general form of the solution in the stratosphere does not change appreciably, only the scale, which depends on h^2, *not* on the sign of $\frac{dU}{dz}$. We may therefore still use Fig. 2 as an approximate description in the more realistic case. In particular we find that the effect of this refinement is a raising of the steering-level. It is a convenient accident that in the case of typical long waves this correction very roughly cancels the correction for variation of ϱ_0 (depression of steering-level) so that we may obtain quite accurate steering velocities by omitting both corrections.

When discussing the system to which Fig. 2 applies we noted the sucking down of the internal boundary in the vicinity of the upper trough (or low pressure centre) and its pushing up in the vicinity of the upper ridge. We may now apply this result directly to the tropopause when there is long wave development. Apart from the agreement with observation so far as the tropopause itself is concerned we may note that we have here a mechanism ("advection" plus "stretching") to account quantitatively for observed changes in ozone measurements by purely dynamical considerations.

IV. Energy Analysis

The above account is based on obtaining complete solutions of the perturbation equations for certain simple initial systems thereby proving the systems unstable and determining the manner of breakdown. It is instructive to consider an alternative analysis in which we merely show the *possibility* of instability.

Nevertheless this more limited analysis makes clear the general nature of the process and makes possible an estimate of maximum growth rate for disturbances (strictly it determines an absolute upper limit to growth-rate) of a wider range of initial systems.

We commence with a system (such as one of those with which we have been concerned) initially in equilibrium and imagine displacements δx, δy, δz at each point, in general functions of x, y, z, t. For simplicity we consider the "incompressible" case (or analogue). The loss in potential energy is computed correct to the second order of small quantities and equated to the gain in kinetic energy, i.e., the kinetic energy associated with a growing disturbance. If we confine our attention to disturbances of "constant shape" in which all the displacements contain the factor $e^{\vartheta_1 t}$ then the kinetic energy involves terms such as $\vartheta_1^2 \cdot \delta\gamma^2$ etc. In this way we obtain ϑ_1^2 as a function of the displacements and, making use of the constraints implied by the continuity equation and one of the momentum equations, we determine an upper limit to ϑ_1^2. If this upper limit is positive we infer that the system is unstable "potentially", i.e., subject to compatibility with other constraints not considered. (In all the cases examined we are able to find physically possible boundary conditions such that this upper limiting growth-rate is actually attained within an arbitrarily small amount. Other boundary conditions yield smaller, but in general comparable, maximum growth-rates. It appears to be a general rule that "potentially" unstable systems are actually so, a feature which may be associated with the infinitude of degrees of freedom of fluid motion.) Apart from the change in sign of ϑ_1^2 the method is essentially the same as that used by RAYLEIGH to determine minimum frequencies of oscillation.

Applied to a barotropic system with uniform velocity this analysis gives an upper limit to the initial rate of growth of a Bénard cell:

$$\vartheta_1^2 \leq -gB \tag{1}$$

Growth-rate is a maximum when δz is everywhere large compared with δx, δy and there is instability only when $B < 0$.

Now consider a baroclinic system. Potential energy is released by a process of "overturning" (we consider the initial stage of this process). It will be convenient to consider separately "vertical overturning" in a plane at right angles to the thermal wind and "quasi-horizontal overturning" in a sloping plane parallel to the thermal wind. The latter case corresponds to the disturbances with which we have hitherto been concerned and it will be convenient to take this case first. Assuming that the disturbances are periodic in the direction of the thermal wind we obtain finally the result:

$$\vartheta_1^2 \leq -g_s \frac{\partial \Phi}{\partial s} \tag{2}$$

if overturning takes place in the xs plane (the s axis being in the yz plane) and g_s is the component of gravity along the s-axis. By hypothesis the isentropic surfaces are not horizontal and if α is their angle of slope (acute angle) we find that ϑ_1^2 is positive if the s-axis slopes at a smaller angle. If α is small we find that growth-rate is a maximum when the s-axis has a slope $\frac{\alpha}{2}$ and then:

$$\vartheta_1^2 \leq \frac{1}{4} g \frac{A^2}{B} \tag{3}$$

and from the definition (II. 12) this is the same as:

$$|\vartheta_1| \leq \frac{1}{2} \cdot \frac{K}{h} \tag{4}$$

Our complete solution (II. 31) therefore corresponds to about 62 % efficiency. The reduction is of course due to the constraints of the rigid boundaries which prevent all the fluid particles being displaced in the optimum direction. (But it is easily verified that in the central region, near $z = 0$, the displacements are nearly in the optimum direction, i.e., along the bisector of the angle between the isentropic surfaces and the horizontal.)

The analogy between (IV. 2) and (IV. 1) is evident — we are here concerned with a kind of "convection" on a large scale, the main displacements being not vertical but in the direction of the s-axis. (Isentropic charts sometimes suggest this kind of picture.)

In the above we have supposed U independent of y as in the systems for which we have obtained complete solutions. We may attempt to generalise the above analysis by considering a system in which initially $V_x = U(z) + W(y)$ but for solutions periodic in the x-direction we have to abandon the assumption of "constant shape". (We are confronted with difficulties similar to those arising in a study of the stability of Couette flow). On the other hand it is easy to study the effect of vertical overturning (in the yz plane) in such a system. We obtain finally:

$$2\vartheta_1^2 \leq -\left[gB - K\left(K - \frac{dW}{dy}\right)\right] + \sqrt{\left[gB - K\left(K - \frac{dW}{dy}\right)\right]^2 - 4\left[(gA)^2 - gB \cdot KK \cdot \left(\frac{dW}{dy}\right)\right]} \tag{5}$$

and if we suppose (as is usually the case):

$$B > 0 : \frac{dW}{dy} < K; \tag{6}$$

the condition for "potential" instability is:

$$(gA)^2 > gB \cdot K\left(K - \frac{dW}{dy}\right) \tag{7}$$

which is the same as:

$$\frac{1}{h^2} > \left(1 - \frac{1}{K}\frac{dW}{dy}\right) \tag{8}$$

There is instability for any Richardson number if $\frac{dW}{dy} > K$, i.e., for an anticyclonic wind shear greater than K, a well-known result but probably not one of very great importance from a practical point of view. On the other hand when $\frac{dW}{dy} \ll K$, (IV. 8) becomes approximately $h^2 < 1$. Thus whereas the atmosphere is normally unstable from the point of view of "quasi-horizontal overturning" it is only so in special circumstances from the point of view of "vertical overturning". We should therefore expect the former process to be dominant in atmospheric development and observation appears to confirm this result at least in middle and high latitudes. (In low latitudes any kind of development is slow unless the static stability is small or negative.) Now the overturning process is an irreversible one and "quasi-horizontal overturning" leads to interchange between warm air at low levels and cold air at high levels. As noted earlier heat is transported *upwards*, statistically balancing radiative cooling at high levels. At the same time heat is normally transported polewards (the main transport is probably associated with long waves) to balance statistically net radiative cooling in high latitudes.

V. The Ultimate Limitations of Weather Forecasting

We may infer from the above analysis (in so far as the atmosphere is always, on a large scale, baroclinic) that atmospheric motion is normally unstable. The fact that practical systems are usually more complicated than those studied does not affect the generality of this result. In fact we have in practice what may legitimately be described as "fully developed turbulence" of a particular kind, the turbulent motion being maintained against frictional dissipation by the growth, from time to time, of disturbances of the kind we have been studying. (The only essential difference between this large-scale turbulence and that occurring on a smaller scale is the manner in which energy is supplied to the turbulent disturbances). Assuming sufficient analytical skill, what are the possibilities of forecasting for such a system? Suppose we attempt to formulate the problem as one of determining a final (forecast) state from a given initial one. The initial state is in practice "given" only within a certain margin of error. For concreteness let us consider pressure at a given point, known within a margin δp. Let ϑ_1 be the maximum growth-rate of unstable disturbances of the system. Then in the final (forecast) state we can guarantee pressure correct only within a margin $\delta p \cdot e^{\vartheta_1 t}$ since disturbances below the margin of error initially (and therefore completely unknown) will have attained this size. It is clear that "guaranteed" forecasts are possible (even in theory) only

for intervals less than t_1 where t_1 is of the order of $\frac{1}{\vartheta_1}$, for beyond this time the margin of uncertainty is so large as to make such "guaranteed" information valueless. Reduction of initial error-margin makes possible only a very limited extension of time interval. For larger time intervals we must reformulate our problem.

Although we cannot, with complete certainty, say anything about long-term developments $\left(t \gg \frac{1}{\vartheta_1}\right)$ it does not follow that all possible developments are equally probable. On the contrary we may infer that probabilities are very unequally distributed (and therefore that information of this kind may be, from a practical point of view, almost as good as "guaranteed" information). As an example consider a set of unstable disturbances of various growth-rates. So long as the determining (perturbation) equations are substantially linear it is clear that the relative importance of the disturbance of maximum growth-rate increases with time i.e., any disturbance composed of components of varying growth-rates will tend towards the size, structure and growth-rate of the "dominant" wave by a process of "natural selection". We may generalise this result by including stable components in the initial perturbation and there can be little doubt that the result is true of *almost* any arbitrary disturbance. Of course in practice conditions are more complicated, the concept of an initial system is less clear-cut and the dominant disturbance is a relatively slowly varying function of the time, quite apart from the modifications which ensue when the disturbance becomes "finite" and the governing equations non-linear. (Moreover we may *choose* to regard certain initial irregularities as "finite" perturbations of a larger system.) Nevertheless the reality of the selection process is made clear every day when we see recognisable, well-known patterns developing "as if from nowhere" which more or less closely resemble in size, structure and behaviour the ideal disturbances we have discussed theoretically.

The above is no more than a prelude to the rather formidable task facing theoretical meteorology — that of discovering the nature of and determining quantitavely all the forecastable regularities of a "permanently unstable" (i.e., permanently turbulent) system. We can be certain that these regularities are necessarily statistical and to this extent our technique must resemble statistical mechanics. But we do not yet know enough about the "atoms" (the life-histories of disturbances) nor are we concerned with "atoms" with a clear-cut individuality. Clearly there are difficulties from the point of view of formulation and it is by no means clear what kind of problem we ought to attempt to solve. But these difficulties are inherent in the study of any kind of turbulent motion and perhaps in the study of irreversible processes (other than isolated ones) in general.

Much of the above formed the subject of a series of colloquia given, at the kind invitation of Profs. J. Bjerknes and C. L. Godske, at the Geofysisk Institutt, Bergen in April, 1947. A more detailed treatment of cyclone theory was given in a doctoral thesis (unpublished: London, 1948). The literature on cyclone and long wave theory is extensive and the writer would like to be excused the compilation of a list of references. He would however like to refer to an independent analysis by J. G. Charney (Journal of Meteorology, Vol. 4, No. 5, Oct., 1947) which in many (but not all) respects is consistent with his own.

The Motion of Barotropic Disturbances in the Upper Troposphere

By George W. Platzman, University of Chicago

(Manuscript received 16 April 1949)

Abstract

Oscillations of a barotropic model of the upper troposphere are studied, taking into account the existence of a zone of intense westerly circulation found in this region of the atmosphere. The fundamental frequencies of these oscillations are determined and some comments are made with regard to their dispersive properties.

Introduction

Since its discovery ten years ago, the planetary type of undulation often prominent in westerly circulations of the troposphere has provided a background against which theoretical and synoptic studies of the middle latitudes have been illuminated in many interesting ways. These undulations occur on a scale which in linear dimensions approximates one-half the planetary radius, while the period of individual 'waves' has an order of magnitude n times the planetary period of rotation, n being the number of waves in a continuous train compassing the hemisphere.

The observations initially providing confirmation of Rossby's (1939) theory were involved mainly in 10,000-ft pressure patterns over the western hemisphere and to a certain extent, particularly in recent years, at somewhat higher elevations. Such "long waves" in the belt of westerlies are found inherently stable, in conformity with the theory, as is evidenced in part by the fact that they are traced frequently through a succession of five to ten days. The interruption of their continuity takes place through the development of smaller-scale disturbances whose operation, although not as yet accessible to satisfactory theoretical analysis, is generally understood to be governed by barocline processes.

The success of the "long-wave" theory and its subsequent evolvements is of particular importance not merely because the theory has provided a satisfactory basis for interpreting certain types of observed motions, but perhaps primarily because it has established so clearly the usefulness of treating a broad class of circulations by means of a dynamical theory in which barotropic, nondivergent, and two-dimensional fluid strata replace the complex structure of the real atmosphere. The kinematical behavior of such a fluid sheet is controlled exclusively by the fact that the vorticity of individual fluid elements cannot be altered. This permanence of vorticity has constituted a central principle which, though achieved through a rigid framework of limitations, is found to extend beyond this framework in its practical usefulness.

The original form of the theory in question is set forth with the help of two simplifications, of which the first is the familiar approximation whereby the zone between two latitude circles not too far separated is supposed mapped on a plane surface so that the language of Cartesian co-ordinates may be used without serious error. The second is the very penetrating device of assuming the latitude variation of the vertical component of the earth's rotation to be uniform throughout the zone, thereby

incorporating a primary feature of the planetary character of the fluid stratum in a simple and effective manner. The equivalent problem involving motions within a truly spherical sheet in which these simplifications are not present was solved subsequently in a well-known investigation by HAURWITZ (1940), although this analysis unavoidably entails complications arising from the necessity of securing cyclic continuity of the motions and from the indeterminacy between different possible modes of oscillation.

In his analysis HAURWITZ retained the full effect of the stabilizing influence of the earth's rotation by postulating a zonal current initially in solid rotation, analogous to the laterally uniform current employed by ROSSBY. Somewhat later, the Cartesian simplifications were reintroduced in a study by GARSTENS (1941) of the coupling between the stabilizing planetary effect and the destabilizing kinematical effect of shearing motion. Recently, a more general inquiry into this problem has been undertaken in an important investigation by KUO (1949).

It has long been recognized that because of the earth's strong gravitational attraction and high rate of rotation, large-scale circulations of the atmosphere are very nearly hydrostatically and geostrophically balanced, and that in consequence, the northward decrease of tropospheric mean temperature must be associated with a predominance of westerly currents in middle latitudes, their intensity increasing until the stratosphere is reached, where the temperature gradient reverses direction. Recently, however, data analyzed by PALMÉN and collaborators have revealed in a striking manner that middle-latitude zonal circulations cannot be accounted for solely by this type of reasoning. It is now known that the intensity and lateral concentration of westerlies in the upper troposphere are in fact so great, particularly during rather regularly recurring periods, as to constitute a distinct and unique phenomenon.

An investigation published by PALMÉN and NEWTON (1948) provides satisfactory material for the purpose of illustrating certain features of the zonal-wind distribution to which attention will be drawn here. In their study, data from twelve synoptic cross sections along the meridian 80° W (all confined to the month December 1946) were averaged by a process in which, at a given pressure level, distance from the polar front was used as a coordinate to establish correspondence between points in the twelve sets of synoptic data. (As a consequence of this procedure it is difficult to assign definite latitudes to the mean cross section but the authors, though cognizant of this awkward feature of their method, have provided such a latitude scale which will be applied in the following discussion without further comment.) In fig. 1 are presented the results of computations based upon the data of PALMÉN and NEWTON[1]; curves A, B, and C show latitude distributions of the zonal wind, absolute angular momentum (per unit mass), and absolute potential vorticity, respectively. These curves were computed from values of zonal geostrophic wind in a quasihorizontal sheet limited above and below by potential temperatures 350° K and 310° K.

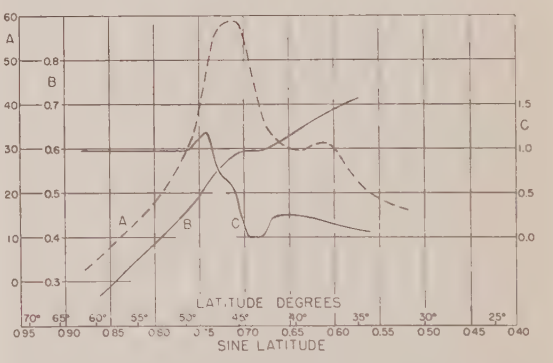


Fig. 1. Latitude distribution of relative zonal wind (curve A), absolute angular momentum (curve B), and absolute potential vorticity (curve C), computed from the semisynoptic data of PALMÉN and NEWTON (1948). The scale for curve A is in meters per second; the scale for curve B is such that the value 1.0 corresponds to the absolute momentum of a unit mass at relative rest on the equator; the scale for curve C is such that the value 1.0 corresponds to the absolute vorticity of the earth at the pole. The method of computation is described in the text.

If Ω_a denote absolute angular velocity of a zonal current, the vertical component of absolute vorticity at latitude Φ when computed along a surface of constant potential temperature (Θ) is

$$Z_\Theta \equiv -\frac{\delta}{\delta\xi}\left[\Omega_a\left(1-\xi^2\right)\right],$$

[1] Through the kindness of the authors the writer had access to the original data.

where δ signifies a variation for which Θ is constant and $\xi \equiv \sin \Phi$ is the sine of the latitude. The corresponding potential vorticity is

$$Z^{\star} \equiv Z_{\Theta} \frac{\partial p_0}{\partial p}$$

In this expression p is the pressure associated with the potential temperature Θ at latitude Φ and p_0 the pressure associated with the same potential temperature at an arbitrarily assigned but fixed latitude Φ_0, while ∂ denotes a variation for which the latitude is constant. In applying this formula to the meridional cross section of PALMÉN and NEWTON, it is convenient and natural to place the latitude at the extreme northern end of their section (58° N) since in this region nearly barotropic conditions prevail; thus, to the north of Φ_0 and also somewhat to the south $\partial p_0/\partial p \approx 1$.

Let $\Theta_1 = 350°$ K and $\Theta_2 = 310°$ K. Curve A in fig. 1 was obtained by computing first a mean value of Ω_a at each latitude from the formula

$$\overline{\Omega}_a \equiv \frac{1}{\Delta p_0} \int_1^2 \Omega_a dp_0,$$

in which Δp_0 is the baric thickness of the layer between Θ_2 and Θ_1 at latitude Φ_0; in the present instance $\Delta p_0 = 100$ mb. The expression $a(\overline{\Omega}_a - \omega) \cos \Phi$ then gives the *relative* linear zonal velocity represented in fig. 1, a and ω being the earth's radius and angular velocity.

Curve B in fig. 1 is obtained from the following nondimensional measure of the absolute angular momentum per unit mass:

$$\overline{M}_a \equiv (\overline{\Omega}_a \cos^2 \Phi)/\omega.$$

The value $\overline{M}_a = 1$ corresponds to the momentum of a unit mass at relative rest on the equator.

Absolute potential vorticity was computed as an average value

$$\frac{1}{\Delta p} \int_1^2 Z^{\star} dp = \frac{1}{\Delta p} \int_1^2 Z_{\Theta} dp_0 = \frac{\Delta p_0}{\Delta p} \overline{Z}_{\Theta},$$

where Δp is the baric thickness of the layer between Θ_2 and Θ_1 at latitude Φ. In curve C the latter quantity is expressed nondimensionally through division by 2ω, the polar value of the earth's vorticity. Reference to the preceding formulae shows that

$$\frac{\Delta p_0}{\Delta p} \frac{\overline{Z}_{\Theta}}{2\omega} = -\frac{1}{2} \frac{\Delta p_0}{\Delta p} \frac{\delta \overline{M}_a}{\delta \xi};$$

from this last expression the distribution pictured in fig. 1 was derived. It should be noted also that the values of Ω_a used in these computations were obtained from the relative angular geostrophic velocity Ω_g by means of the gradient wind formula

$$\Omega_a = \omega (1 + 2\Omega_g/\omega)^{1/2},$$

which provides for the effect of curvature in zonal flow.

The curves illustrate several features of the wind distribution that have been noted in previous investigations. Roughly 250 km south of the relative wind maximum is a narrow zone of uniform absolute momentum and zero absolute vorticity, while approximately the same distance north of the relative wind maximum the potential vorticity attains a maximum value which in this case is about 20 per cent in excess of the very uniform values prevailing to the north. It is of course difficult to assign this maximum value with certainty because curve C is essentially a second derivate of the basic data, i.e., of the isobaric contour heights, but the reliability of the estimate is perhaps somewhat enhanced by the method of averaging through a relatively deep layer as described in the paragraphs preceding.

Similar computations were performed utilizing mean data published recently by HESS (1948); the results are shown in fig. 2. These data, also along the meridional plane 80° W, comprise eight months of radiosonde records, January and February 1942—45, for twelve stations located between latitudes 3° S and 74° N. The curves in fig. 2 were based on the following choice of parameters: $\Phi_0 = 70°$ N, $\Theta_1 = 375°$ K, $\Theta_2 = 325°$ K; again in this case $\Delta p_0 = 100$ mb. As would be expected, the features so apparent in the semisynoptic data assembled by PALMÉN and NEWTON are much less prominent here, particularly the rather abrupt transition from high values of absolute vorticity on the north side of the current axis to low values on the south, as shown by curve C in fig. 1.

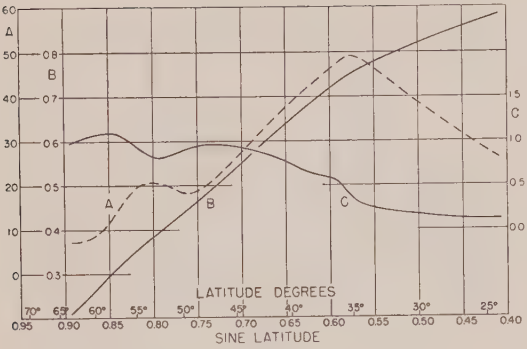


Fig. 2. Latitude distribution of relative zonal wind (curve A), absolute angular momentum (curve B), and absolute potential vorticity (curve C), computed from the mean data of Hess (1948). For further description see legend under fig. 1.

The vorticity distribution pictured in fig. 1 serves as a guide in constructing an appropriate distribution of zonal velocity for the purpose of investigating two-dimensional barotropic oscillations of the upper troposphere. From an inspection of this distribution it is apparent that there exist three relatively distinct zones, each with a more or less characteristic type of vorticity variation: a northern zone of uniformly high vorticity, a transition zone in which the vorticity decreases from a high to a low value, and a southern zone of low vorticity. In order to simplify the mathematical analysis, only the northern and southern zones have been incorporated in the theoretical model, which is illustrated by the dotted line in fig. 3. It is evident that the transition zone

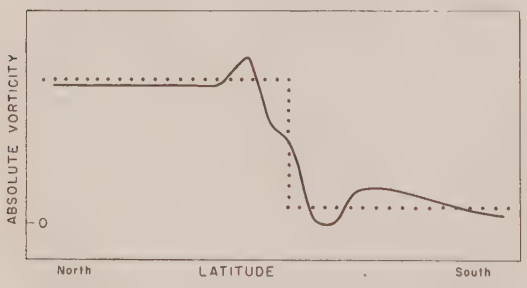


Fig. 3. Schematic hypothetical distribution of absolute vorticity (dotted line) compared with the distribution computed from the data of Palmén and Newton described in fig. 1.

is replaced by a vorticity discontinuity coinciding with the axis of the relative zonal current, while to the north and to the south the vorticity is independent of latitude.

Theoretical velocity distribution of the mean motion

Fixing attention upon a level surface in the upper troposphere, north of a given latitude the zonal flow in this level is characterized by uniform absolute vorticity, which for simplicity is assumed to prevail throughout a polar cap whose southern boundary coincides with the axis of the zonal current. The zonal velocities corresponding to this condition are easily calculated.

Denoting by Ω_a the absolute angular velocity of the zonal current in latitude Φ, the total circumpolar circulation of the absolute linear velocity will be

$$2\pi a^2 \Omega_a \cos^2 \Phi,$$

where a is sensibly the mean radius of the earth. This circulation must evidently be equal to the surface integral of the absolute vorticity, taken over the polar cap north of latitude Φ. Since the absolute vorticity is assumed uniform, the integral in question has the value

$$2\pi a^2 Z (1 - \sin \Phi),$$

where Z is the value of the absolute vorticity. On comparing the two expressions it is found that[1]

$$\Omega_a \cos^2 \Phi = Z (1 - \sin \Phi);$$

the corresponding relative angular velocity is $\Omega = \Omega_a - \omega$, where ω is the earth's angular velocity.

At the latitude Φ_0 of the current axis[2] the absolute vorticity passes abruptly from the value Z prevailing in the polar cap to the smaller value $Z - \Delta Z$, denoting by ΔZ the positive value of the discontinuous change of vorticity. At a latitude Φ, south of Φ_0, the total circumpolar circulation of the absolute linear velocity is, as before,

$$2\pi a^2 \Omega_a \cos^2 \Phi,$$

while, assuming continuity in *velocity* at

[1] This angular-velocity distribution was first given by Rossby (1947), who postulated the existence of a polar cap of uniform vorticity on theoretical grounds.

[2] The symbol Φ_0 as used henceforth has no connection with the same symbol involved previously in discussion of figs. 1 and 2.

latitude Φ_0, the absolute circulation at Φ_0 is given by

$$2\pi a^2 Z (1 - \sin \Phi_0).$$

The difference between these two expressions must be equal to the surface integral of $Z - \Delta Z$ taken over the zone between Φ and Φ_0. Since $Z - \Delta Z$ is supposed uniform, this integral has the value

$$2\pi a^2 (Z - \Delta Z) (\sin \Phi_0 - \sin \Phi);$$

hence, to the south of Φ_0,

$$\Omega_a \cos^2 \Phi =$$
$$= Z (1 - \sin \Phi) - \Delta Z (\sin \Phi_0 - \sin \Phi).$$

If symmetry is required between the northern and southern hemispheres, the southern-hemisphere polar cap will possess the uniform vorticity $-Z$; further, the absolute vorticity $Z - \Delta Z$ in the equatorial zone between latitudes Φ_0 and $-\Phi_0$ will, on the present hypothesis, be uniformly zero, so that $\Delta Z = Z$. The preceding equation then reduces to

$$\Omega_a \cos^2 \Phi = Z (1 - \sin \Phi_0),$$

which corresponds to uniform absolute angular momentum in the equatorial zone.

Barotropic disturbances in the upper troposphere

Consider a horizontal stratum of the atmosphere in which the motion is purely zonal and possesses the angular-velocity distribution described in the preceding paragraphs. Neglecting the effects of heterogeneity, suppose a disturbance to be initiated in this layer, the velocities being purely horizontal and nondivergent. Under these conditions the absolute vorticity of each fluid element is conserved. If the disturbance is initiated without the injection of additional vorticity, it follows that the initially uniform distribution of absolute vorticity cannot be altered.[1] Therefore, a dynamically possible oscillation of a steady current possessing an initially uniform distribution of absolute vorticity may be obtained by superposition of an *arbitrary irrotational disturbance*, subject only to such restrictions as may be imposed by boundary conditions.[1]

In the absence of eddy stresses the dynamical equations for horizontal motion may be expressed in the form

$$\frac{du}{dt} - \left(2\omega \sin \Phi + \frac{u}{a} \tan \Phi\right) v = -\alpha \frac{\partial p}{\partial x}$$

$$\frac{dv}{dt} + \left(2\omega \sin \Phi + \frac{u}{a} \tan \Phi\right) u = -\alpha \frac{\partial p}{\partial y}$$

The time is denoted by t; u and v are the eastward and northward velocity components relative to the earth; Φ is the latitude and ω the earth's angular velocity; a is the mean radius of the level surface, or effectively, the earth's mean radius; α is the specific volume; p is the pressure. In this and subsequent notation the variations $\partial/\partial x$ and $\partial/\partial y$ are not intended as symbols for differential coefficiencs, but should be interpreted as directional derivatives; formally, $\partial x = a \cos \Phi \, \partial\lambda$ and $\partial y = a \partial\Phi$, where λ is the longitude angle. The equations are thus essentially in spherical coordinates.

By expanding the individual derivative the preceding equations can be arranged in the form

$$\frac{\partial u}{\partial t} - Zv = -\frac{\partial}{\partial x}\left(\alpha p + \frac{1}{2} q^2\right)$$

$$\frac{\partial v}{\partial t} + Zu = -\frac{\partial}{\partial y}\left(\alpha p + \frac{1}{2} q^2\right),$$

where $q^2 \equiv u^2 + v^2$ and

$$Z = 2\omega \sin \Phi + \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} + \frac{u}{a} \tan \Phi$$

is the absolute vorticity. This arrangement involves the assumption that the fluid stratum

[1] The distribution is uniform apart from the discontinuous change from the value Z to the value $Z - \Delta Z$ at the current axis. However, in studying the disturbance we neglect possible effects of lateral mixing, and assume that the discontinuity will not diffuse into a finite zone.

[1] It is illuminating to draw upon the analogy to autobarotropy in stratified fluids, which is established when the "equation of piezotropy" (corresponding here to the conservation of absolute vorticity) is equivalent to the "equation of barotropy" (corresponding here to the state of uniform vorticity).

is homogeneous (α uniform). On linearization[1] the preceding equations reduce to

$$\left.\begin{aligned} \frac{\partial u}{\partial t} - Zv &= -\frac{\partial}{\partial x}(\alpha p + Uu) \\ \frac{\partial v}{\partial t} + Zu &= -\frac{\partial}{\partial y}(\alpha p + Uu), \end{aligned}\right\} \quad (1)$$

in which u and v are now understood to be the components of a small irrotational disturbance, p being the perturbation pressure, while U is the relative linear velocity of the undisturbed zonal motion with uniform absolute vorticity Z.

Since the perturbation motion is irrotational, there exist a stream function ψ and velocity potential φ such that

$$u = -\frac{\partial \psi}{\partial y} = -\frac{\partial \varphi}{\partial x}; \quad v = +\frac{\partial \psi}{\partial x} = -\frac{\partial \psi}{\partial y};$$

and if in addition to these relations it is noted that Z is uniform in time and, except for certain specified discontinuities, in space, the dynamical equations are found to possess the first integral

$$\alpha p + Uu = Z\psi + \partial\varphi/\partial t,$$

in which an arbitrary function of the time arising in the integration is, as customary, supposed merged in φ. Denoting by Ω the relative angular velocity of the undisturbed zonal current, so that $U = a\Omega \cos \Phi$, the preceding 'Bernoulli' equation may be expressed in the form

$$\alpha p = Z\psi + \frac{\partial \varphi}{\partial t} + \Omega \frac{\partial \varphi}{\partial \lambda},$$

where λ, as noted earlier, is the longitude angle in spherical coordinates.

The boundary conditions to be satisfied along the line of vorticity discontinuity at latitude Φ_0 may be formulated in a simple manner. The disturbance, being small, deforms the line of discontinuity only slightly from its original location along a latitude circle. Hence the kinematical boundary condition, which requires continuity of normal velocity across this line, is equivalent to the continuity of v, the northward velocity component; thus $\Delta v_0 = 0$, where Δ is used to signify a difference in value at essentially the same point but on opposite sides of the discontinuity, and the subscript is used to indicate values taken at the latitude Φ_0. Since $v = \partial\psi/\partial x$ it is clear that to the first order of small quantities the condition $\Delta v_0 = 0$ is equivalent to

$$\Delta \psi_0 = 0. \quad (2)$$

The dynamical boundary condition, requiring continuity of pressure, may be obtained directly from the Bernoulli equation; thus, since $\Delta p_0 = 0$ and $\Delta \psi_0 = 0$,

$$\psi_0 \Delta Z + \Delta \frac{\partial \varphi_0}{\partial t} + \Omega_0 \Delta \frac{\partial \varphi_0}{\partial \lambda} = 0,$$

provided also that the undisturbed zonal current velocity is continuous ($\Delta\Omega_0 = 0$). It will be helpful at a later stage to introduce here the *convective* coordinates

$$l \equiv \lambda - \Omega_0 t, \quad \tau \equiv t,$$

by means of which an observer moving downstream at the angular rate Ω_0 would measure the disturbance. In terms of these coordinates the preceding equation takes the form

$$\nu \psi_0 + \tfrac{1}{2} \Delta \frac{\partial \varphi_0}{\partial \tau} = 0, \quad (3)$$

where $\nu \equiv \frac{1}{2} \Delta Z$.

It remains now to establish the proper solution of $\nabla^2 \psi = 0$ or $\nabla^2 \varphi = 0$ compatible with the boundary conditions (2) and (3). To represent the two-dimensional solution of Laplace's equation appropriate for a spherical surface it is convenient to note that this solution may, by conformal transformation, be expressed in polar coordinates, by stereographic projection of the spherical surface onto the polar plane. The stereographic (polar) coordinates are r, λ where

$$r \equiv \sec \Phi - \tan \Phi,$$

and λ is the longitude. In polar coordinates

[1] More generally, the problem considered may be treated by methods known in connection with Poisson's equation, the distribution of vorticity being assigned over a two-dimensional surface. The following more restricted approach has been preferred as being simpler for the purpose in view.

the characteristic solution of Laplace's equation is $r^n e^{in\lambda}$.

North of latitude Φ_0 and also south of this latitude, the complex potential $w = \varphi + i\psi$ can be represented generally by a sum of elementary independent harmonics w_n:

$$w = \sum_{n=1}^{\infty} w_n.$$

It is evident that each harmonic w_n must satisfy the boundary conditions (2) and (3), since these conditions are linear. The appropriate form of w_n, regarded as a function of r and of the convective coordinates l and τ, may be written as follows:

$$w_n = a_n r^n e^{inl} \qquad (r < r_0)$$

$$\overline{w}_n = \overline{b}_n r^{-n} e^{-inl}, \qquad (r > r_0)$$

where the dependence of the complex coefficients a_n and $\overline{b}_n$ on t (or τ) is to be determined in accordance with (3). In these equations w_n applies to the region north of Φ_0 and vanishes at the north pole ($r = 0$), while $\overline{w}_n$ applies to the region south of Φ_0 and vanishes at the south pole ($r = \infty$). Since $\overline{\psi}_{n0} = \psi_{n0}$ according to (2), it follows that $\overline{b}_n = -a_n^{\star} r_0^{2n}$ (the asterisk designating a complex conjugate), and from this, $\varphi_{n0} = -\varphi_{n0}$. Therefore $\Delta\varphi_{n0} = 2\varphi_{n0}$, so that the dynamical boundary condition (3) is equivalent to

$$\partial\varphi_{n0}/\partial\tau + \nu\psi_{n0} = 0.$$

Thus the real part of the complex function $\partial w_n/\partial\tau - i\nu w_n$ vanishes on the circle $r = r_0$ within which the function is analytic. This function must then vanish everywhere:

$$\partial w_n/\partial\tau - i\nu w_n = 0.$$

Assembling these results, the solution compatible with (2) and (3) finally takes the form

$$w_n = c_n r^n e^{i(nl+\nu\tau)} \qquad (r < r_0)$$

$$\overline{w}_n = -(r/r_0)^{-2n} w_n^{\star}, \qquad (r > r_0)$$

where c_n is an absolute (complex) constant, the value of which is determined by the initial form of the disturbance, as yet unspecified. The general solution, obtained by superposition of all elementary harmonics w_n, is

$$w = e^{i\nu\tau} \sum_{n=1}^{\infty} c_n r^n e^{inl}, \qquad (r < r_0)$$

together with a similar expression for the region $r > r_0$.

It is evident by inspection of the form of this solution that with respect to the 'convective' observer ($l = \lambda - \Omega_0 t = \text{constant}$) the disturbance is *periodic*, each configuration of the motion being repeated after a time interval $2\pi/\nu = 4\pi/\Delta Z$, where ΔZ is the discontinuity of vorticity in the axis of the current. (If ΔZ is expressed as $\Delta Z = \varkappa\omega$, this period has the value $2 \div \varkappa$ in days.) No systematic dispersion can occur in such disturbances. This is also evident by inspection of the form of the elementary harmonics w_n; for, writing

$$\mu_n = \Omega_0 - \nu/n, \qquad (4)$$

it is found that in terms of λ and t,

$$w_n = c_n r^n \exp\,[in(\lambda - \mu_n t)].$$

The periodic wave train represented by the nth harmonic is therefore propagated eastward, without change of structure, at the angular rate μ_n given by (4). With respect to the convective observer this wave train is propagated westward at the angular rate ν/n and hence in the time interval $t = 0$ to $t = 2\pi/\nu$ (independent of n) has traveled an angular distance $2\pi/n$ equal to one wave cycle. Superposition of all harmonics must then produce the same disturbance as existed for $t = 0$.

Equation (4) is the fundamental frequency equation for disturbances in a zonal current for which the absolute vorticity is uniform to the north and to the south of a latitude where the relative velocity of the current reaches a maximum value. The vorticity to the north exceeds that to the south by an amount ΔZ. In equation (4), μ_n is the angular velocity of an elementary wave train (positive when the waves move eastward) and Ω_0 is the angular velocity in the axis of the zonal current, while $\nu = \frac{1}{2}\Delta Z$ and n is the total number of in-

dividual waves encompassing the hemisphere.[2]

Rossby (1947) has pointed out that a regime of uniform absolute vorticity within an extensive polar cap would tend to be established through the operation of large-scale lateral mixing, particularly during periods of fully developed meridional exchange by cyclonic and anticyclonic eddies. Further, he has suggested that a limiting value of the vorticity of such a polar cap would be given by 2ω, the polar vorticity of the earth. If $Z = 2\omega$ and if, for example, the absolute vorticity $Z - \Delta Z = 0$ south of the current axis, then $\Delta Z = Z = 2\omega$ so that $\nu \equiv \frac{1}{2}\Delta Z = \omega$. With these values of Z and ν it is possible to obtain a simple expression for the number of waves (hemispheric wave number) in a stationary simple harmonic wave train. Setting $\mu_n = 0$ and $\nu = \omega$ in (4), the stationary wave number is given by

$$n_s = \omega/\Omega_0.$$

If the value of Ω_0 is calculated from the angular-velocity distribution described earlier, then, provided $Z = 2\omega$, it is easily found that

$$n_s' = \frac{1 - \sin \Phi_0}{1 + \sin \Phi_0}.$$

For $n_s = 3$ this gives $\Phi_0 = 30°$; for $n_s = 4$, $\Phi_0 \approx 37°$; for $n_s = 5$, $\Phi_0 \approx 42°$. The latter values appear to be in fair conformity to conditions frequently prevailing in the upper troposphere, as observed on isobaric contour charts.

Dr George Cressman has drawn the writer's attention to his studies of long-wave patterns at the 500-mb level, which indicate that in most cases southward displacement of the latitude of maximum wind is accompanied by an *increase* in hemispheric wave number. However, the purpose of the above formula relating n_s and Φ_0 is merely to indicate that the present theory provides reasonable correspondence between plausible values of hemispheric wave number and the latitude of maximum wind. Without making any special assumption as to the value of ν in (4), the stationary wave number would be

$$n_s = \nu/\Omega_0.$$

It is the writer's impression that with southward displacement, there is a concurrent intensification of the maximum wind and of the difference in shear north and south — that is, southward displacement is accompanied by an increase in both ν and Ω_0. Cressman's results would then be consistent with the relation $n_s = \nu/\Omega_0$, provided ν increases more rapidly than Ω_0.

The analysis leading to equation (4) is defective in at least one obvious respect. It was tacitly assumed that south of latitude Φ_0 the solution $\bar{w}_n$ extends to the south pole, whereas if this is the case, the main zonal motion, corresponding to $Z - \Delta Z = \text{constant}$, would also extend to the south pole, giving infinite velocities there. If symmetry is required between the two hemispheres, the frequency equation (4) is altered somewhat. To demonstrate this, assume that south of latitude $-\Phi_0$ in the southern hemisphere, the uniform vorticity $-Z$ prevails, while in the intermediate (equatorial) zone the absolute vorticity is zero. The discontinuity in vorticity is then $\Delta Z = Z$; this applies to both discontinuities, at Φ_0 in the northern hemisphere and at $-\Phi_0$ in the southern hemisphere.

If transverse velocities do not appear across the equator ($\bar{\psi} = 0$ for $r = 1$), it suffices to restrict consideration to the form of the disturbance in the northern hemisphere, since the southern-hemisphere disturbance will follow from symmetry. The appropriate form of w_n, regarded as a function of r and of the convective coordinates l and τ, may be expressed as:

$$w_n = a_n r^n e^{inl}$$
$$\bar{w}_n = \bar{a}_n r^n e^{inl} + \bar{b}_n r^{-n} e^{-inl},$$

the former applying to the polar cap north of Φ_0 ($0 < r < r_0$) and the latter to the symmetrical equatorial zone between Φ_0 and $-\Phi_0$ ($r_0 < r < r_0^{-1}$). The dependence of the complex coefficients a_n, $\bar{a}_n$, and $\bar{b}_n$ on τ is to be determined as before, by means of (3). From the symmetry condition $\bar{\psi}_n = 0$ for

[2] In his early studies of the stability of sensitive jets, Rayleigh (*Theory of sound*, chapter 21) obtained frequency equations similar in form to (4); indeed, the physical model of a jet employed by Rayleigh is, apart from obvious adaptations to the geophysical setting of the present discussion, essentially equivalent to the model from which (4) has resulted.

$r = 1$, and from the condition (2) that $\psi_n = \bar{\psi}_n$ for $r = r_0$, it is found that

$$\bar{a}_n = -a_n r_0^{2n}/\sigma_n; \quad \bar{b}_n = -a_n^{\star} r_0^{2n}/\sigma_n,$$

where $\sigma_n \equiv 1 - r_0^{2n}$. With these values for the coefficients of $\bar{w}_n$, it follows that

$$\Delta\varphi_{n0} \equiv \varphi_0 - \bar{\varphi}_{n0} = 2\varphi_{n0}/\sigma_n.$$

The dynamical boundary condition (3) is then equivalent to

$$\partial\varphi_{n0}/\partial\tau + \nu_n \psi_{n0} = 0,$$

in which $\nu_n \equiv \sigma_n \nu = (1 - r_0^{2n})\nu$. Thus the real part of the complex function $\partial w_n/\partial\tau - i\nu_n w_n$ vanishes on the circle $r = r_0$ within which the function is analytic. This function must then vanish everywhere:

$$\partial w_n/\partial\tau - i\nu_n w_n = 0.$$

It follows that the solution providing symmetry between the hemispheres finally takes the form

$$w_n = c_n r^n \exp[i(nl + \nu_n\tau)]$$
$$\bar{w}_n = -(r/r_0)^{-2n}(w_n^{\star} + r^{2n} w_n)/\sigma_n,$$

in which c_n, as before, is an absolute constant, the value of which is determined by the initial form of the disturbance.

The frequency equation in this case is evidently

$$\mu_n = \Omega_0 - \nu_n/n. \qquad (5)$$

This differs from (4) in that ν is replaced by $\nu_n = (1 - r_0^{2n})\nu$, but since $r_0^2 < 1/3$ for $\Phi_0 > 30°$, equation (5) will give values of μ_n differing only slightly from (4) except for the very lowest harmonics, $n < 4$.

It is possible that equation (4) may prove useful prognostically as a means for estimating the displacement of long-wave patterns in the upper troposphere, during periods when such patterns are well developed. In applying this formula, a profile of the zonal wind must first be constructed and plotted, for example, in meters per second versus degrees latitude. The displacement may then be computed by measuring the following quantities:

1. The latitude Φ_0 of the maximum zonal wind,
2. The maximum zonal wind speed U_0,
3. The difference ΔZ between the mean absolute vorticity to the north and to the south of the maximum wind,
4. The wave length or hemispheric wave number n estimated from analysis of 500- or 300-mb contours.

From these four quantities the rate of displacement of the wave pattern in degrees longitude per day may be determined from (4) or appropriate modifications thereof. Preliminary tests of this formula in a few cases indicate that it is capable of providing correct orders of magnitude of the wave displacements.

Energy propagation

Considerable attention in recent years has been directed toward the dispersive properties of "planetary" disturbances in connection with the growth and decay of the tropospheric westerlies and in relation to the coupling between widely separated zones of activity. This interest has been aroused mainly by the provocative investigation published about four years ago by Rossby (1945). As a consequence of this and later studies by the same author, it has become evident that the purely barotropic component of all large-scale circulations may perform an important function by operating as channels through which local concentrations of atmospheric energy are, in part, permitted to travel outwards in advance of the parent disturbance and activate latent supplies of energy successively in different longitudes, thus serving, in this sense, as perpetual catalytic agents. Many aspects of the dispersion problem have been examined critically in a study published recently by Yeh (1949).

Rossby's analysis has involved in large measure the well-known correspondence between the so-called kinematical and dynamical group velocities. If μ_n is the elementary wave speed (in angular measure) and n the corresponding wave number, then from a *purely formal* point of view the kinematical group velocity would be expressed by $d(n\mu_n)/dn$. But it must be observed that this expression cannot be interpreted strictly in accordance with the ordinary interference theory of wave groups, since the elementary components into which an arbitrary phase distribution on the sphere can ultimately be decomposed are necessarily of integral order, and thus constitute a discrete rather than a continuous spectrum of wave numbers. Further, since the present dis-

cussion concerns disturbances in which harmonics of low order are predominant, this difficulty is apparently not easily disposed. In dealing with wave-like disturbances it is, of course, possible to define a phase angle π so that an expression such as $\cos \pi$, combined with an appropriate amplitude factor, adequately represents the disturbance. The phase velocity would then be given by

$$-\partial\pi/\partial t \div \partial\pi/\partial\lambda,$$

but there is no reason to suppose that the two quantities forming this ratio represent, respectively, the local frequency and local wave number, assuming the possibility of defining these local wave properties kinematically. This identification between phase gradients and kinematically relevant wave properties is limited evidently to those portions of the disturbance where these gradients are sufficiently small, measured in terms of time and space units comparable to the local period and wave length.

It has been customary in hydrodynamical studies, following particularly the work of RAYLEIGH, to consider the dynamical group velocity as a rate of propagation of energy. Since the analytical properties are known for the particular type of disturbance investigated in the preceding section of this paper, it is possible to compute directly the rate of propagation of wave energy for each elementary component of the disturbance, although such a computation does not in itself lead necessarily to information regarding the presence or absence of dispersion, since the energy distribution in a simple harmonic wave train is always uniform. The procedure in question is based upon the following line of reasoning.

Consider a fixed meridian F and a moving meridian M (fig. 4) with angular rate of displacement γ. Let $E_t(\lambda, t)$ represent the rate of transmission of wave energy through the fixed meridian between the pole and an arbitrary latitude Φ and let $E_\lambda(\lambda, t)$ represent the "density" of wave energy in this range of latitude, i.e., $E_\lambda\,\delta\lambda$ is the amount of wave energy in the space between the pole, latitude Φ and two meridians separated by an angle $\delta\lambda$. At the instant when M coincides with F the rate of transmission of energy through M is

$$E_t(\lambda, t) - \gamma E_\lambda(\lambda, t).$$

If the total flux of energy through the moving meridian is made to vanish after a time interval T during which this meridian has been displaced from λ_0 to $\lambda_0 + \gamma T$, then

$$\int_0^T E_t(\lambda_0 + \gamma t, t)\,dt = \gamma \int_0^T E_\lambda(\lambda_0 + \gamma t, t)dt.$$

For periodic waves propagated without change of form at the angular rate μ, every wave property is such that, for example, $E_t(\lambda, t) = E_t(\lambda - \mu t)$; hence

$$\int_0^T E_t(\lambda_0 - \gamma' t)\,dt = \gamma \int_0^T E_\lambda(\lambda_0 - \gamma' t)\,dt,$$

where $\gamma' \equiv \mu - \gamma$. Writing $\xi \equiv \lambda_0 - \gamma' t$,

$$\int_{\lambda_0}^{\lambda_0 - \gamma' T} E_t(\xi)\,d\xi = \gamma \int_{\lambda_0}^{\lambda_0 - \gamma' T} E_\lambda(\xi)\,d\xi.$$

The moving meridian will traverse one complete wave if $\gamma' T$ is equal to the length of the individual waves, in which case the choice of λ_0 is immaterial in view of the periodicity of the wave train. It follows that

$$\gamma = \overline{E_t}/\overline{E_\lambda},$$

the bar signifying an average value taken over one complete wave cycle. This formula is equivalent to that employed by RAYLEIGH (1877). On the average, there is no net flux of energy through a plane moving uniformly with speed γ, which in this sense therefore represents the rate of propagation of wave energy.

To determine the appropriate expressions for E_t and E_λ it is convenient to return to the perturbation dynamical equations (1). On multiplying the first of these by u, the second by v, and adding, it is found that

$$\frac{\partial}{\partial t}\left(\frac{1}{2}q^2\right) + \operatorname{div}(\chi u, \chi v) = 0,$$

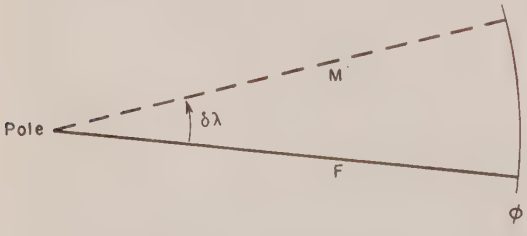


Fig. 4. Schematic diagram illustrating fixed meridan (F) and moving meridian (M) used in computation of energy velocity.

where $\chi \equiv \alpha p + Uu$ and $\frac{1}{2} q^2 \equiv \frac{1}{2} (u^2 + v^2)$ is the perturbation kinetic energy (per unit mass). This is one form of the continuity equation for wave energy. Upon integration over a fixed area and transformation of the divergence integral into a line integral, the equation expresses that the rate at which perturbation energy within the area is increasing, must equal the rate at which energy is transmitted across the boundaries of the area. The perturbation energy is transmitted in part by the perturbation pressure p and in part advectively by the basic current U. In spherical coordinates,

$$\frac{\partial}{\partial t}\left(\frac{1}{2} q^2\right) + \frac{1}{a \cos \Phi}\left(\frac{\partial \chi u}{\partial \lambda} + \frac{\partial \chi v \cos \Phi}{\partial \Phi}\right) = 0.$$

The wave energy in an elementary region $\delta\lambda$, $\delta\Phi$ is $\frac{1}{2} q^2 a^2 \cos \Phi \, \delta\Phi \delta\lambda$ (apart from certain constant factors), so the energy "density" defined previously is

$$E_\lambda = \int_{\Phi_0}^{\pi/2} \frac{1}{2} q^2 a^2 \cos \Phi \; \delta\Phi.$$

Further, it is evident from the form of the energy continuity equation that $\chi u a \delta\Phi$ expresses the rate at which wave energy is transmitted in the direction of wave propagation, normal to the segment $\delta\Phi$, so that

$$E_t = \int_{\Phi_0}^{\pi/2} \chi u a \delta\Phi.$$

The latitude Φ_0 in these expressions for E_t and E_λ will be taken as that of the maximum wind, where the regime of uniform vorticity Z is interrupted. On multiplication of the preceding energy continuity equation by $a^2 \cos \Phi \, \delta\Phi$ and integration from Φ_0 to $\pi/2$,

$$\frac{\partial E_\lambda}{\partial t} + \frac{\partial E_t}{\partial \lambda} = \chi_0 v_0 a \cos \Phi_0.$$

The right-hand member of this form of the continuity equation determines the rate at which energy is accumulated or depleted due to lateral transfer across the latitude Φ_0. In the case of a periodic wave train this must on the average be equal to zero since the wave does not propagate laterally, as is easily verified by noting that χ and v are 90 degrees out of phase.

From the Bernoulli equation obtained previously,

$$\chi \equiv \alpha p + Uu = Z\psi + \partial\varphi/\partial t,$$

where φ and ψ are the perturbation velocity potential and stream function. It may be shown without difficulty that for the periodic wave train progressing at the rate μ_n without change of form, in which $\psi \sim r^n$,

$$Z\psi + \partial\varphi/\partial t = (\mu_n + Z/n)\, au \cos \Phi.$$

It follows that

$$\bar{E}_t = \left(\mu_n + \frac{Z}{n}\right) \int_{\Phi_0}^{\pi/2} \overline{u^2} a^2 \cos \Phi \delta\Phi =$$

$$= \left(\mu_n + \frac{Z}{n}\right) \bar{E}_\lambda,$$

since $\overline{u^2} = \overline{v^2}$ in this case. Hence the rate of propagation of energy through a simple harmonic wave train in the region of uniform vorticity Z is given by

$$\gamma_n = \mu_n + Z/n \qquad (6)$$

In this expression, either of the formulas (4) or (5) for the wave speed μ_n may be substituted, depending upon whether a symmetrical disturbance is required in the southern hemisphere. But whatever the form of the disturbance south of the vorticity discontinuity, the dynamical group velocity γ_n computed here exceeds the wave speed by an amount Z/n, which increases with increasing wave length.

It cannot be concluded, however, that an arbitrary disturbance, composed of an infinite number of such wave trains, is necessarily dispersive. In fact, it has already been pointed out that disturbances for which equation (4) gives the harmonic frequencies are nondispersive, because of the linearity of this equation. Thus, the energy distribution associated with disturbances of this kind, must be propagated without change. To demonstrate this, it is convenient to refer to the relation

$$\partial w_n/\partial\tau = ivw_n$$

obtained in the preceding section. Here w_n is the complex potential for the nth harmonic and $v \equiv \frac{1}{2} \Delta Z$. Since v in this case is inde-

pendent of n, the total potential $w = \Sigma w_n$ must satisfy the same condition:

$$\partial w/\partial \tau = ivw.$$

Differentiate this equation with respect to λ and separate real and imaginary parts:

$$\partial u/\partial \tau = +\nu v; \quad \partial v/\partial \tau = -\nu u.$$

From these it is evident that

$$\frac{\partial}{\partial \tau}\left(\frac{1}{2}q^2\right) = \left(\frac{\partial}{\partial t} + \Omega_0 \frac{\partial}{\partial \lambda}\right)\left(\frac{1}{2}q^2\right) = 0,$$

the first equality resulting merely from the meaning of $\partial/\partial\tau$. The energy distribution of an arbitrary disturbance is therefore propagated convectively without change. It is interesting to note, however, that at every latitude $\Phi > \Phi_0$, where the convective velocity Ω is smaller than Ω_0, the disturbance energy is propagated at a rate $\Omega_0 > \Omega$.

In the more realistic model for which the harmonic frequencies are given by (5), the situation is quite different; in this case ν is replaced by $\nu_n = (1 - r_0^{2n})\,\nu$, and an arbitrary disturbance must, in general, be dispersive. The spreading of an initial concentration of energy may, of course, be computed from the appropriate FOURIER series, but unfortunately a simple means of describing the character of the dispersion has not been found.

Acknowledgments. — I wish to express my appreciation for the benefit of many fruitful discussions with Prof. V. P. STARR, and Dr. T. C. YEH.

REFERENCES

GARSTENS, M. A., 1941: Some effects of shearing motion on the propagation of waves in the prevailing westerlies. *Papers Phy. Ocean. Meteor., Mass. Inst. Tech. & Woods Hole Ocean. Instn*, **8**, no. 4, 32 pp.

HAURWITZ, B., 1940: The motion of atmospheric disturbances on the spherical earth. *Journal of Marine Research*, **3**, 254—267.

HESS, SEYMOUR L., 1948: Some new mean meridional cross sections through the atmosphere. *Journ. of Meteor.*, **5**, 293—300.

KUO, HSIAO-LAN, 1949: Dynamic instability of two-dimensional nondivergent flow in a barotropic atmosphere. *Journ. of Meteor.*, **6**., 105—122.

PALMÉN, E., and NEWTON, C. W., 1948: A study of the mean wind and temperature distribution in the vicinity of the polar front in winter. *Journ. of Meteor.*, **5**, 220—226.

RAYLEIGH, LORD, 1877: On progressive waves. *Proc. London Math. Soc.* (1), **9**, 21—26. (Reprinted in *Theory of Sound*, Appendix to vol. 1.)

ROSSBY, C.-G., 1939: Relation between variations in the intensity of the zonal circulation of the atmosphere and the displacements of the semi-permanent centers of action. *Journal of Marine Research*, **2**, 38—55.

— 1945: On the propagation of frequencies and energy in certain types of oceanic and atmospheric waves. *Journ. of Meteor.*, **2**, 187—204.

— 1947: On the distribution of angular velocity in gaseous envelopes under the influence of large-scale horizontal mixing processes. *Bull. Amer. Meteor. Soc.*, **28**, 53—68.

YEH, TU-CHENG, 1949: On energy dispersion in the atmosphere. *Journ. of Meteor.*, **6**, 1—16.

NOTE

Survey of the Activities of the Radiation Commissions of the International Meteorological Organisation and of the International Union of Geodesy and Geophysics

By ANDERS ÅNGSTRÖM, Swedish Meteorological and Hydrological Institute

In many scientific activities the war caused an interval of stagnation. We are now faced not only with the problem of taking up and reestablishing activities interrupted several years ago, but also with the problem of reorganizing these activities on a new basis with due regard to present needs.

With these conditions in mind an historic survey of the scope and the activities of the Radiation Commissions may be of interest.

The Solar Radiation Commission of the I.U.G.G. was established at the General Assembly of the Union at Madrid in the year 1924 with H. H. Kimball as its first president and with A. Ångström, Gorczynski, Maurain, Platania and Simpson as members. According to instructions its first activities were to be directed toward promoting pyrheliometric and other specified observations. From the report of Dr Kimball to the General Assembly at Prag in 1927 it may be seen that the work of the Commission in the three year interval between the meetings in Madrid and Prag was concentrated on the organization of a number of stations for recording total radiation from sun and sky, namely in Belgian Congo, in Samoa, on Spitzbergen and at Tromsö. In a paper in Monthly Weather Review for April 1927, submitted to the Meteorological Association, Kimball had given a program for the treatment of pyrheliometric measurements and records in order to get from them results of general significance.

At Prag the Solar Radiation Commission[1] was given more definite instructions and was authorized to:

(1) encourage, by all possible means, the maintenance of an international net of stations for measuring solar radiation,

(2) which ought to include as many high altitude stations as possible,

(3) take steps to insure that especially the stations at high altitudes would occupy themselves with measurements of ultraviolet radiation and of the ozone content of the atmosphere.

In a special resolution the Commission was authorized to cooperate with the Solar Radiation Commission of the International Meteorological Organization (IMO) in order to organize comparisons between pyrheliometric standard instruments used in different countries, to prepare a program and a practical instruction for daily observations and to facilitate the speedy publication of unusual results.

At the General Assembly in Stockholm in 1930 a number of papers were presented dealing with the subjects indicated above, particularly with the distribution of radiation over the earth's surface and with the ultra-violet radiation. On the initiative of Dr Dobson, and with special regard to his investigations, a subcommission of the Solar Radiation Commission was established for studying the ozone problem. This subcommission, consisting of Messrs Abbot, Ångström, Chalonge, Dobson, Fabry, Götz, Kimball and Ladenburg, was to pursue researches on atmospheric ozone mainly according to the outlines drawn up by Dr Dobson at the meeting at Stockholm. At this meeting Dr Kimball

[1] In the Minutes from Prag the Commission is for the first time referred to as the Solar Radiation Commission.

resigned as president and was succeeded by the author of this survey.

Under the circumstances prevailing in 1930 it was regarded as an important task to insure close collaboration between the two Commissions. The IMO Commission had been established as early as 1896 in Paris. Its scope had never been very clearly defined and it had been appreciably broadened with due regard to the fact that various new radiation problems of meteorological interest had, in the course of time, been brought into the field of scientific activity. Thus in 1919 a commission of a somewhat more astrophysical character instituted in Southport 1903 by the International Union for Solar Research had been fused into the IMO Commission and its field of work seems to have covered as well the solar radiation in general as also various radiation processes within the atmosphere. It may be assumed, however, that its primary aim must have been to organize radiation measurements within the meteorological net works belonging to the meteorological institutes. But it must be admitted that this aim was very often regarded a secondary one and that with some justification the scientific aspects also by this Commission were considered of primary importance. Consequently, to great extent, the fields of action of the two Commissions were then actually overlapping. There were two means by which duplication of work could be avoided, either to make a clear division of responsibilities of the two Commissions, so that no overlapping would occur, or to collaborate so closely that they actually would function as one single Commission.

At that time one of the most important questions within actinometric research was undoubtedly the one concerning the comparability of results obtained in different parts of the world with different actinometers. The difference between the two prevailing scales for solar radiation, namely the Ångström scale and that of the Smithsonian Institution, was still in 1930 as large as between 3 and 5 per cent, having amounted to 9.2 per cent in 1908 and to 5.5 per cent in 1911, before some sources of error in the Smithsonian scale had been corrected. In 1914 I pointed out and investigated the existence of a small source of error of about 1.5 per cent in the Ångström compensation pyrheliometer and this result was later confirmed through an investigation by MARTEN. However, if this correction was applied, there remained nevertheless a difference between the two scales of between 2.5 and 3.5 per cent. There were strong reasons for believing this difference not to be constant but varying with the degree of turbidity of the atmosphere and consequently also with the height above sea level, and possibly also with other factors not clearly known at that time.

For all questions concerning comparisons of actinometric measurements, touching as they do upon the possible variability of the sun and upon the possible existence of long period fluctuations in solar radiation, it seemed of fundamental importance

(1) to establish, through experiments and field measurements the real difference between the two prevailing scales, and its dependence on various meteorological factors,

(2) to establish, on the basis of a study of the errors adhering to the present scales, a true pyrheliometric scale.

This fundamental investigation falling within the field of activity given to both Commissions, it was regarded appropriate that a very close collaboration between them was established. Since 1930 the two Commissions have in fact functioned very nearly as a joint Commission of the International Meteorological Organization and of the International Association of Meteorology, this being promoted through the arrangement that practically all members of the one Commission were also members or were elected members of the other one.

The work on the pyrheliometric standard scale was thus carried on in common by both Commissions and it was to a considerable extent promoted by grants from the I.U.G.G.

Asubcommission "pour la pyrhéliométrie absolue" was instituted in Frankfurt a.M. in 1932 by the IMO Commission with Professor Süring as president and Messrs Abbot, Ångström, K. Fuessner, O. Hoelper, W. Mörikofer and F. F. Volochine as members, Mörikofer being secretary. A meeting of this subcommission was held in Potsdam in February 1935. The Minutes from this meeting, at which much of the subsequent work was organized, is given as an appendix to the Minutes of the meeting of the two Radiation Commissions at Oxford 1936, where a number of rather important reports on standard pyrheliometry were given by Süring, Mörikofer and by Guild of the National Physical Laboratory, London.

The results reached at the time of the outbreak of the war were not final, but they were, as far as they had gone, highly instructive as regards the problem in view. On the basis of the results at present available, we may give the following table as

regards the relative values of the two scales, the Ångström scale being put equal to 100 %

Ångström (uncorrected)	100.0 %
„ corrected for edge effect, according to Ångström and Marten	101.3 %
Smithsonian Scale 1913	103.5 %
Solar Radiation Commissions 1936	101.0 %[1]
Guild, Physical Laboratory 1934	101.3 %
Abbot (Smithsonian scale 1934)	101.2 %

The results of the Radiation Commissions, of Guild and of Abbot consequently agree within 0.3 per cent. The mean value 101.2 agrees further within 0.1 per cent with the value of the Ångström scale, corrected with 1.3 per cent for the edge effect pointed out by myself. The common Ångström scale generally in use, which is not corrected for the edge effect, is according to these results 1.2 per cent too low, the Smithsonian Scale of 1913, 2.3 per cent too high. It may be added that the results of the Astrophysical Observatory of the Smithsonian Institution are still, for the purpose of continuity evidently given in the scale of 1913.

It must here be emphasized, however, that these results as far as Ångström pyrheliometers and Silverdisc pyrheliometers (Smithsonian scale) are concerned refer mainly to mean values obtained at low level stations.[2] As the readings of the two types of pyrheliometers are to different extent influenced by the diffuse sky light entering on the side of the sun within the opening angle of the pyrheliometer, the difference between the two scales, *as represented by the pyrheliometers in actual use*, are to a certain degree dependent upon the state of the sky. Thus Mörikofer has found a difference between the readings of Ångström and Silver disc instruments amounting to 5—6 per cent at high level stations, at about 1,600 m above sea level. At very clear sky the difference is consequently larger than at more hazy sky and it will probably be possible to express the difference as a function of air mass and turbidity.

Further investigations have shown that unless certain precautions are taken some minor sources of error are liable to affect as well the reading of the Ångström pyrheliometers as the Silverdiscs. Dr Curvoisier at the observatory of Davos is now, with a grant from the Swedish National Board of Scientific Research, occupied with a detailed study of this matter. His investigations, which to large extent have been planned in collaboration with the president and secretary of the Radiation Commission, are now devoted to the purpose of building a *permanent standard pyrheliometer* founded on the principle of electric compensation, with avoidance of sources of error adhering to older types. He proposes also, on the basis of researches already to large extent carried out, an accurate determination of all secondary effects, such as convection, reflection power of black surfaces, etc., which arise in the measuring device.

On the side of these studies, the Radiation Commission was rather early brought to investigate the question of the *turbidity* of the atmosphere. The matter was brought before the Commission through several reports and papers especially by Linke and Ångström, later by Hoelper, Kimball and Götz. The Commission stimulated these works through special meetings and discussions. The clear definitions of the "Trübungsfaktor" of Linke and of the "turbidity coefficient" of Ångström have been of value for illuminating the turbidity condtions of various air masses at different altitudes and latitudes, and comparative studies of the conditions of turbidity at different parts of the world have been made possible.

Another question taken up by the Commission refers to the standardization of glass filters used in actinometric measurements for studying different parts of the spectrum. The Commission promoted the introduction of certain standard types, stimulated exact measurements and information concerning their transmission, through the Meteorological institute at Potsdam it arranged to keep in store considerable blocks of glass of certain given properties for use in glass filters delivered on demand. This work was interrupted and seems partly to be in disorder after the war.

The Commission has promoted the work on an international bibliography on actinometry. This task was as a consequence of a resolution carried at the meeting at Stockholm in 1930 and was undertaken by M. Volochine. The work has been supported through several grants from the Union. The bibliographical cards at present distributed cover the time 1900—1929 (without abstracts) and 1930—35 (with abstracts). As several meteorological and geophysical journals now have the publishing of abstracts on their program, a continuation of this work by the Commission seems scarcely to be justified.

[1] 2.2 per cent lower than Smithsonian scale.

[2] A special report on the Davos's studies on standard pyrheliometry will be given by Dr W. Mörikofer. See especially Mörikofer 1939, pp. 4056—57.

As a consequence of an application (1936) of the Radiation Commission to the President of the Association of Oceanography, a small grant was given to a subcommission of the "Conseil permanent pour l'Exploration de la Mer" to carry out investigations of methods for submarine light-measurements and of the best methods for standardization of intruments involved. The work was carried out in intimate collaboration with representatives of the Radiation Commission and the results are published in a couple of papers by Johnson and Olsson,[1] in which methods for calibration and use of photoelectric cells for the said purpose are described.

As may be seen from the survey given above the activity of the Radiation Commission has been concentrated on problems of a fundamental nature, namely *standardization and development of instruments and methods*. The rather unsatisfactory state of conditions prevailing concerning instruments and methods in the field of radiation, and the necessity to bring about an improvement in this regard, has formed a justification for treating more superficially such questions as the organization of a net work of actinometric stations and promoting measurements at high levels and measurements in the ultraviolet. Before such measurements are organized in large scale, it seems necessary to require that the instruments and methods used must give results which are strictly comparable. To furnish a more solid basis in that respect has been the chief aim of the Commission.

Having thus given a summary of the work of the Solar Radiation Commission in the past, I finally propose to express in a few words my opinion concerning the development in the future.

It seems evident that the work now in progress for the purposes of establishing a standard pyrheliometric scale and creating real standard instruments which can be kept at one or more observatories or institutes as normals, that this work must be continued and pursued vigorously through international cooperation. Some kind of permanent supervision and guidance with regard to the matter of instruments and methods in the field of radiation is also necessary.

The entire question has to some extent come in a new position through a recent action by the International Meteorological Organization. At the meeting in London 1946, the first after the end of the war, the IMO Conference of Directors took up the question of reorganizing the body, especially with regard to its commissions. In London the Solar Radiation Commission was, together with a number of scientific commissions, not reestablished. Some of the duties of the old Solar Radiation Commission were instead laid upon a new Commission on Actinometry, instituted in Toronto with myself as president, as a subcommission under the *Commission for Instruments and Methods*, which already had been established in London.

The instructions given by the Head commission included for the Subcommission on Actinometry the following duties:

(a) To recommend the most suitable equipment for measurements of solar and atmospheric radiation at meteorological stations.

(b) To insure that radiation measurements in different countries be made on a uniform and immediately comparable basis by arranging for the establishment of international standard scales and for inter-comparison of substandards.

(c) A continued investigation in order to establish pyrheliometric standard scales and investigations of errors adhering to radiation instruments in current use.

(d) An investigation of the sensitivity of different sunshine recorders in current use. It is desirable that standard colorimetric specifications for the cards used with the Campbell-Stokes Sunshine Recorder be introduced.

(e) To prepare a terminological list on various names entering into work on radiation, especially as regards instruments.

(f) To prepare instruction for the use and standardization of radiation instruments in current use.

(g) To make recommendations, in collaboration with the Climatological Commission, on questions of time, observations, and form of publication concerning data on radiation.

For my part I consider the step taken by the IMO to mean a development toward a clearer and more effective organization, *provided* that some of the duties of the old Solar Radiation Commission which are not covered by the Commission on Actinometry are cared for by some other international body. Such a body exists already in the International Union for Geodesy and Geophysics. Through the establishment of the Radiation Commission of its Meteorological Association, an international body

[1] Johnson and Olsson (see appended list of Minutes and Reports). The result of this research have been of great use, especially during the Albatross expedition under the direction of Prof. Hans Pettersson.

is already in existence with principal duty to cover the scientific field of the problems of radiation.

The action taken by the International Meteorological Organization invites to a clear separation between the more technical side covering instruments and methods within actinometry and the more scientific one, dealing with the scientific problems and their solution.

For the technical part the Commission on Actinometry is the right forum. It ought to concentrate especially on instruments and methods which are suitable for stations within the meteorological network.

The more scientific part falls very naturally under the Meteorological Association of the I.U.G.G., the work of which according to its instruction "should primarily lie in the discussion of scientific subjects". (See Resolution IV of the Meteorological Section at its meeting in Stockholm 1930.)

The scientific problems with which the commission of the Union has to deal cover at present a rather wide field of research. Investigations of the variations of the solar constant in various parts of the solar spectre, studies of the ultraviolett radiation and its variation with various factors, high altitude investigations of outgoing and incoming radiation, the ozon layer and its influence on the radiation balance, may be especially mentioned here. It seems clear, however, that the work of the commission of the Union for promoting the study of these and related questions must consist chiefly in a stimulation of interest through discussion and addresses and through personal contacts, than through an organization and administration involving the introduction of routine methods.

It is evident to my mind that great precaution must be taken in the attempts to obtain general acceptance of routine methods and standard instruments in fields where the scientific basis must still be worked out. In cases where the scientist himself, occupied with the solution of a scientific problem asks for help through international administration, such a help ought certainly be given after due consideration, but the desires to organize ought not to press themselves upon the scientific activity.

LITERATURE

List on reports contained in the Minutes of the International Radiation Commissions:

1. Procès-verbaux des séances de la section de Météorologie. Union Géodesique et Géophysique Internationale, Deuxième Assemblée Generale à Madrid — Octobre 1924, Rome 1925.
2. Procès-verbaux (de la même section) Troisième Ass. Gen. à Prague — Septembre 1927, Rome 1928.
3. Procès-verbaux (de la mème section) Quatrième Ass. Gen. à Stockholm Août 1930 (1) Actes de la section (2) annexes, Paris 1933.
4. Procès-verbaux de l'association de météorologie. Union Géodesique et Géophysique Internationale, Sixième Ass. Gen. Edinbourg — Septembre 1936, Paris 1937.
5. Report of the President of the International Commission for Solar Radiation, Meeting at Copenhagen, 1929, Utrecht 1930.
6. Protokolle der Sitzungen der Internationalen Strahlungskommission in Frankfurt a/M — September 1932, Leyde 1933.
7. Protokolle der Sitzungen der Internationalen Strahlungskommissionen in Oxford 1936, Leyde 1937.

In the following list the reports are marked with letters, which refer to the meetings at which the report was presented, and in the Minutes of which the report consequently is printed. The letters (M), (P), (S), (L) and (E) indicate thus that the report is printed in the minutes of the Meteorological Association of the International Union for Geodesy and Geophysics from Madrid (M), Prag (P), Stockholm (S), Lisboa (L) and Edinburgh (E), the letters (C), (F) and (O), that the report is printed in the Minutes of the Solar Radiation Commission of the International Meteorological Organization from Copenhagen (C), Frankfurt (F) and Oxford (O).

SHAW, SIR NAPIER: Radiation in relation to meteorology. Summary of the present state of our knowledge (M).

AUSTIN, E. E.: Tables of the results of observations of solar radiation at 59 stations in various parts of the world (M).

GORCZYNSKI, L.: Sur la nécessité d'exploration des regions caractéristiques de la terre, au point de vue de la répartition spectrale du rayonnement solaire en relation avec les conditions météorologiques (M).

GORCZYNSKI, L.: Sur les nouveaux procédés thermoélectriques pour l'étude de la radiation solaire (M).

GORCZYNSKI, L.: Sur l'état actuel des mesures pyrhéliométriques et spectrales du rayonnement solaire (M).

GORCZYNSKI, L.: Suggestions relatives à l'exploration spectrale des régions caractéristiques de la Terre (M).

MAURAIN, CH.: Mesure de la radiation solaire à l'observatoire du Parc St-Maur, près Paris (M).

RICHARDSON, L. F.: Survey of the reflectivity of the earth's surface as seen from aeroplanes (M).

KIMBALL, H. H.: Report of the special commission appointed by the Meteorological Section of the inter-

national union of Geodesy and Geophysics at its meeting in Madrid in October 1924 to promote pyrheliometric and other specified observations (P)

MELANDER, PALAZZO, OISHI, MERCANTON, CURRY, ÅNGSTRÖM, A.: Organisation des observations de la radiation solaire dans les différent pays (P).

ÅNGSTRÖM, A.: On the unit of radiation used in meteorological treatises on actinometry (P).

CHAPMAN, S.: Report of the committee appointed by the international research council to further the study of solar and terrestrial relationships (P).

1933, Report of the Commission on the relation of the meteorological section of the International Union of Geodesy and Geophysics with the International Meteorological Committee (S).

KIMBALL, H. H.: Commission de radiation solaire. Report (S).

DOBSON, G. M. B.: Research on atmospheric Ozone (S).

KIMBALL, H. H.: Instructions for obtaining solar radiation intensity measurements (S).

ÅNGSTRÖM, A.: On the atmospheric transmission of sun radiation (S).

VOLOCHINE, F.: Les travaux du laboratoire actinométrique international de Trappes (S).

VOLOCHINE, F.: Le bulletin actinométrique international de l'observatoire de Trappes (S).

GORCZYNSKI, L.: Quelques résultats des mesures actinométriques en mer effectuées dans les oceans Atlantique et Indien de 1923 à 1928 (S).

KALITIN, N. N.: Contribution à l'étude de l'intensité de la radiation de la voute céleste au voisinage du soleil (S)

BRAZIER, C. E.: Rapport sur les travaux actinométriques à l'observatoire de Parc Saint-Maur 1927—1930 (S).

LINDHOLM, F.: Densité de diffusion optiques caractérisant l'atmosphère comme milieu trouble (S).

BOREEMA, J.: Ultraviolet solar radiation on Java (S).

ÅNGSTRÖM, A.: Rapport de la Commission de radiation solaire (L).

FABRY, CH.: Les progrès des Methodes pour l'étude de l'ozone atmospherique (L).

GORCZYNSKI, L.: Sur l'utilité du dosage de la vapeur d'eau dans l'atmosphère terrestre par la méthode spectrographique (L).

ÅNGSTRÖM, A.: Rapport de la Commission de Radiation Solaire (E).

ÅNGSTRÖM et MÖRIKOFER 1937: Lettre adressée par les Commissions de Radiation Solaire de l'O.M.I. et de l'U.G.G.I. au Président de l'Association de Météorologie de l'U.G.G.I. (R).

WEHRLÉ, PH.: Rapport sur la Bibliographie Actinometrique (E).

ABBOT, C. G.: Observations of the solar constant of radiation (C).

ÅKERBLOM, F.: Konstantenbestimmung an Pyrgeometern (C).

ÅNGSTRÖM, A.: Plan for investigations of atmospheric radiation (C).

DORNO, C. und LINDHOLM, F.: Zur Standard-Pyrheliometer-Frage (C).

DORNO, C. und LINDHOLM, F.: Zur Cadmiumzellen-Frage (C).

HERGESELL, H.: Strahlungsmessungen in Flugzeugen (C).

BÜTTNER, K.: Bericht über Messungen der Sonnen- und Himmelsstrahlung im Flugzeug (C).

LINKE, F.: Durchlässigkeit von Farbfiltern (C).

KALITIN, N. N.: Les travaux actinométriques dans l'U.R.S.S. (C).

KALITIN, N. N.: Les travaux actinométriques à l'observatoire magnétique et météorologique de Pavlovsk (Sloutzk) U.R.S.S. (C).

STENZ, EDWARD: Compte rendu des travaux sur la radiation solaire effectuées en Pologne en 1926—28 (C).

SÜRING, R.: Bericht über die aktinometrischen Arbeiten am Meteorologischen Observatorium Potsdam (C).

SÜRING, R.: Das Potsdamer Pyrheliometer für Absolutmessungen (F).

VOLOCHINE, F. E.: Quelques appareils du Laboratoire actinométrique à Trappes (F).

ABBOT, C. G.: Briefliche Mitteilung (F).

MÖRIKOFER, W.: Neue Vergleichungen zwischen dem Silver-disk-Pyrheliometer und dem Ångström-Pyrheliometer (F).

LUNELUND, H.: Bericht über Untersuchungen der Sonnen- und Himmelsstrahlung in Finnland 1929—32 (F).

LINKE, F.: Bericht über den Stand der Trübungsforschung (F).

FEUSSNER, K.: Die Schott'schen Farbmessgläser OG 1 und RG 2 (Potsdamer Normalfilter), ihre physikalischen Konstanten und ihre Verwendung in der meteorologischen Strahlungsforschung (F).

ÅNGSTRÖM, A.: Ueber die Trübung der Atmosphäre (F).

GÖTZ, F. W. P.: Trübung und anomale Extinktion (F).

VON DEM BORNE, H.: Experimente mit künstlicher Trübung (F).

ÅNGSTRÖM, A.: Bericht über Ausstrahlungsmessungen (F).

FEUSSNER, K.: Beitrag zur Pyrgeometerfrage (F).

MÖLLER, F. und MÜGGE, R.: Ein Auswertungspapier für gradlinige Wärmestrahlung (F).

GÖTZ, F. W. P.: Ueber den heutigen Stand der Ozonfrage (F).

MÖRIKOFER, W.: Erfahrungen über die Vergleichbarkeit von Cadmiumzellenmessungen (F).

LINKE, F.: Ueber das Oeffnungsverhältnis von Aktinometern (F).

KALITIN, N. N.: De la radiation de la voute céleste autour du soleil pour les ondes longues et pour les ondes courtes (F).

STENZ, E.: Bericht über die Untersuchungen der Wasserdampfabsorptionsbänder im ultraroten Sonnenspektrum (F).

Sitzungsprotokoll und Richtlinien der Subkommission für die Organisation der strahlungsklimatischen Forschung (F).

SÜRING, R.: Die Feussner'schen Versuche mit dem Potsdamer Absolutpyrheliometer (O).

MÖRIKOFER, W.: Ueber die zwischen Kompensations- und Silverdisk-Pyrheliometern beobachteten Differenzen (O).

GUILD, J.: Notes on the scale of radiation intensity and the calibration of secondary radiometers for solar observations (O).

HOELPER, O.: Trübungsbestimmung durch Filtermessungen (O).

GÖTZ, F. W. P.: Zur Trübungsfrage (O).

LINKE, F.: Erfahrungen mit dem Trübungsfaktor (O).

ÅNGSTRÖM, A.: On the coefficient of turbidity (O).

KIMBALL, H. H.: Letter regarding turbidity-coefficient (O).

JENSEN, CHR.: Ueber die Wichtigkeit von Polarisationsmessungen (O).
LINKE, F.: Registrierungen des Sonnenspektrums (O).
LIANA, F.: Advantages of recording potentiometers for solarigraphic use (O).
KALITIN, N. N.: Des erreurs pendant le mesurage de l'intensité de la radiation solaire, importées par les nuages dans l'angle d'ouverture de l'actinomètre (O).
ABBOT, C. G.: The Melikeron — a nocturnal radiation instrument (O).
HOELPER, O. und MÖRIKOFER, W.: Zur Frage der Standardisierung der Cadmiumzellen (O).
GÖTZ, F. W. P.: Bericht über das UV-Dosimeter (O).
PETTERSSON, H. and ÅNGSTRÖM, A.: Photo-Elements for Submarine Light-Measurements (O).
MÖRIKOFER, W.: Erfahrungen mit dem Aktinographen Robitzsch (O).
COBLENTZ, W. W.: The Evaluation of Ultraviolet Solar Radiation in Absolute Units (O).
STENZ, E.: Ueber Strahlungsmessungen in den Wigry-Seen (O).

Other literature referred to in the Survey:

ÅNGSTRÖM, A.: 1914: Das Ångströmsche Kompensationspyrheliometer und die pyrheliometrische Skala. Meteorologische Zeitschrift, Heft 8, pp 5.

MARTEN, W., 1922: Zur Frage der absoluten pyrheliometrischen Skala. Meteor. Zeitschrift, Heft 10, pp. 3.

MÖRIKOFER, W., 1939: Meteorologische Strahlungsmethoden. Handbuch der biologischen Arbeitsmethoden, herausgegeben von Dr E. Abderhalden, Abt. II, Teil 3, pp. 240.

JOHNSON, N. G. and OLSSON, H. 1943: On the standardization of photoelectric elements by means of solar radiation. The total energy of incident radiation computed from records with photoelectric elements.

(These lastnamed papers are contained as well in Communications: Series of Papers No. 47 of the Swedish Meteorological and Hydrological Institute and also in: Geografiska Annaler, 1943, H. 1—2, pp. 24.)